# FROM DEVIANCE TO MEANING: FRONTAL TEMPORAL INTERACTION IN THE PROCESSING OF UNEXPECTED EVENTS

BY

CHUN-YU TSE

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Psychology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Doctoral Committee:

Professor Gabriele Gratton, Chair
Professor Monica Fabiani, Director of Research
Professor Gary S. Dell
Professor Greg A. Miller
Associate Professor Susan M. Garnsey

**Abstract**

Deviance detection refers to the ability to detect changes from an expected regularity. This ability allows us to monitor and react to the environment when the unexpected happens. Extensive laboratory studies have been conducted to understand the underlying neural mechanisms and cognitive processes in deviance detection, however, the functional roles of brain regions involved in deviance detection remain unclear. The current thesis aims at studying the spatiotemporal dynamics of brain activity in deviance detection by measuring Event-related Optical Signal (EROS). EROS is a relatively new brain imaging method which has relatively good spatial and temporal resolutions to allow the spatiotemporal dissociation of brain responses. Specifically, three experiments were conducted to study the interactions of the frontal and temporal cortices in 1) detecting semantic and syntactic violations, 2) representing regularities, and 3) detecting audiovisual deviance. The first experiment investigated brain responses to semantic and syntactic violations in sentence comprehension. Similar temporal followed by frontal cortex activities were elicited by both semantically and syntactically anomalous words. However, the temporal activity corresponding to a semantic anomaly was more ventral than that corresponding to a syntactic anomaly. The second experiment investigated the brain response to the counterpart of deviance detection – regularity detection. Sequences of auditory tones governed by three contingency rules were presented. Temporal and fronto-parietal network activities were observed according to the processing requirements of the contingency rules. This result suggests that the brain can simultaneously hold different models of the stimulus contingency within the information processing stream, but that these representations are held at different levels, both in terms of latency and location of the brain responses. The last experiment extended the investigation of brain responses in deviance detection from unimodal to multimodal sensory systems. By using a set of control conditions, brain responses to audiovisual deviance

detection were separated from those of audiovisual integration. More interestingly, interactions of audiovisual integration and audiovisual deviance detection were revealed. The results from these experiments and previous EROS studies suggest that deviance detection is a common property among various cognitive processes and involves similar basic cognitive components in the frontal and temporal cortices.

*This dissertation is dedicated to the memory of W.-C. Tse*

**Acknowledgements**

I could not have completed my Ph.D. training without the help and support from the following:

No words can fully express my greatest gratitude to my Ph.D. advisors, Monica Fabiani and Gabriele Gratton, for their guidance and support in the past five years. I am sure what I learnt from them will continue to be valuable to my future career. I would also like to extend my gratitude to my dissertation committee members, Gary S. Dell, Susan M. Garnsey, and Greg A. Miller for providing valuable suggestions and advice.

My plan for Ph. D. study would not be materialized without the encouragement and advice from my master's thesis advisor Trevor Penney. I am incredibly indebted to his continuous support and advice for my career and personal life.

I am grateful to my colleagues in Monica and Gabriele's lab for sharing the sweet and bitter life of graduate study, including Kathy Low, Ed Maclin, Brian Gordon, Nils Schneider-Garces, James Lee, Mark Fletcher, Chandramalika Basak, and Tanya Stanley. Equally important are the lab alumni, Echo Leaver, Elena Rykhlevskaia, Eunsam Shin, Carrie Brumback-Peltz, Christopher Whalen, Emily Wee, and Jeff Sable, who provided survival tips in the American culture, graduate school, and the lab. Special thanks to previous lab staff, Yukie Lee, Lupe Arroyo, Jayme Jones, Brian Martinez and Jonathan Kimnach, and collaborators, Charlene Lee, Jennifer Black, Jaimie Gilbert, Michael Novak, Teemu Rinne, and Eric K. K. Ng. This dissertation was supported by the NIMH, ONR, Beckman Institute, Carle hospital, Beckman Institute graduate student fellowship and UIUC psychology department fellowship.

Finally, I would like to dedicate this thesis to my mother, sister, my sister's fiancé, Mandy Choi, and Arare for their unconditional support, care and consideration.

## Table of Contents

# Chapter 1

## General Introduction

Deviance detection refers to the ability to detect changes from an expected regularity. This ability allows us to monitor and react to the environment when the unexpected happens. Deviance detection processes are only triggered when an unexpected event happens, allowing us to allocate our limited attention capacity efficiently. Extensive laboratory studies have been conducted to understand the underlying neural mechanisms and cognitive processes in deviance detection. Other than being studied as a research topic, deviance detection has also been used as a tool to study other cognitive processes, e.g., perception, memory system, language, and response conflict.

As summarized in a detailed review by Fabiani (2006), deviance detection can take place at different stages along the information processing stream, from perception, sensory memory, working memory to semantic and syntactic violation detection in language. Recent brain imaging studies focused on identifying the neural substrates involved in various types of deviance detection. Other than localizing the brain regions involved, understanding the temporal dynamics of different brain regions in deviance detection is equally important. The current thesis aims at investigating the spatiotemporal dynamics of the brain in deviance detection with Event-related Optical Signal (EROS). Specifically, we are interested in the interactions of the frontal and temporal cortices in 1) semantic and syntactic violation, 2) regularity representation, and 3) audiovisual deviance detection.

Because event-related potentials (ERPs) have been widely used in studying deviance detection, a review of some of the most studied ERP components related to deviance detection

will be presented and the involvement of the temporal and frontal cortex will be discussed. Next, how EROS can help in the understanding of the temporal dynamics between the temporal and the frontal cortices will be explained. In the current work I will focus on auditory and verbal processing although deviance detection may also occur in other modalities (Czigler, 2007; Kekoni, et al., 1997; Kimura, Katayama, & Murohashi, 2005; Kimura, Katayama, Ohira, & Schröger, 2009).

**N1**

Auditory N100, or N1, is an ERP component that refers to the negative deflection at around 100 ms to150 ms after the onset of the change (Näätänen & Picton, 1987; Picton, Alain, Otten, Ritter, & Achim, 2000). Basically, any change in the physical characteristics of stimuli that associates with energy change can elicit an N1. In other words, the onset or offset of a tone, or the change in frequency or intensity within a continuous tone all elicit an N1, although these stimulation changes might not necessarily be related to perceptual changes (Picton, Alain, Otten, Ritter, & Achim, 2000). N1 also tracks the temporal characteristic of a train of stimuli. The amplitude of N1 is inversely proportional to the repetition rate of the stimuli, and this modulation effect is related to the refractory effect of the N1 generators (Budd, Barry, Gordon, Rennie, & Michie, 1998), or top-down inhibitory processes (Sable, Low, Maclin, Fabiani, & Gratton, 2004). In any case, the amplitude of N1 is sensitive to the violation of temporal regularity in the auditory domain.

The auditory N1 typically shows a frontal central maximum with mastoid inversion when using nose reference. This scalp distribution suggests that the N1 originates from the superior temporal plane. MEG studies (Pantev, et al., 1995; Williamson & Kaufman, 1981) revealed the

topographic distribution of the magnetic counterpart of the N1 response in the auditory cortex. Guiraud et al. (2007) further demonstrated that the topographic distribution of N1 can be used to monitor plastic changes in the auditory cortex in deaf people after cochlear implant. Specifically, they showed that the N1 dipole orientation was modulated by the frequency of the auditory stimuli after cochlear implant, indicating the restoration of topographic organization in auditory cortex.

**MMN**

Pre-attentive change of an auditory stimulus elicits another ERP component called Mismatch Negativity (MMN), which is independent of the N1described above. MMN is typically investigated with the passive oddball paradigm, in which a series comprising deviant tones embedded in a train of standard tones is presented to the participant, while the participant reads a book or watches a silent subtitled movie. The deviant tones could be different from the standard tones in any feature, e.g., pitch, intensity, duration (Näätänen, 1992), and more interestingly, violates regularities established by the standard tones, i.e. rules governing the inter-tone relationship (Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001). MMN is defined as a negative deflection in the difference waveform with the standard subtracted from the deviant (Näätänen & Michie, 1979; Näätänen, Paavilainen, Rinne, & Alho, 2007). It has a frontocentral distribution and usually peaks at around 100-200ms. Similar to the N1, MMN also showed mastoid inversion with the nose electrode as the reference (Näätänen, 2007; Näätänen & Winkler, 1999). Different from the N1, the amplitude of the MMN corresponds to the perceived degree of deviance, and is associated with the just noticeable difference and the subjective/perceptual difference in an active discrimination task (Näätänen, 1992).

MMN was originally hypothesized to reflect a sensory memory process (Näätänen, Jacobsen, & Winkler, 2005). More specifically, MMN is not the sensory memory itself, but reflects the mismatch that arises from the comparison of the memory trace between the standard and deviant. For example, in a typical passive oddball paradigm, the frequent standard tones establish a mnemonic template containing all of the physical characteristics of the standard tone, and the new incoming tone is compared against this standard template. If there is any difference in the comparison with incoming tones, i.e. the deviant tones, MMN is elicited (Näätänen, 2001).

The characteristics of MMN are consistent with the construct of echoic memory (a form of sensory memory), which was proposed to be a modality specific process holding large amount of information with high fidelity for a brief period of time without requiring attention (Atkinson & Shiffrin, 1968). However, recent studies have shown that the MMN responds not only to the differences in the physical or perceived characteristics of the tones, but also to the violation of the regularity established by the standard. The MMN was further proposed to be a detector for regularity or prediction violation (Schröger, 2007; Winkler, 2007).

Under this model, MMN is elicited when there is a violation to the regularity or the abstract rule set up by the standards. The incoming tones are analyzed and separated into different auditory sources or streams. Sound features are extracted to establish a representation for the auditory object. Regularities among the auditory object are also detected and integrated into the model of the acoustic environment. This model is used to predict future auditory events. When a new incoming tone is perceived by the listener, the analysis procedure described above is repeated and the representation of the incoming tone is compared against the predictive model. If there is a violation of the regularity or the prediction, MMN is elicited to update the predictive model or to reduce the weight of the deviant tones to be integrated into the current prediction

model. Thus, the predictive model is dynamically updated for every incoming sound. The traditional sensory memory mismatch model can be treated as a special case of the regularity violation model, when the regularity is established by physical characteristics stored in sensory memory. However, in the regularity violation model, the establishment of the "standard" only requires regularity not repetition of identical sounds.

Evidence supporting the regularity violation model come from various studies using abstract rules or non-repeating standard tones. Regular presentation time, i.e. SOA, of the standard tones establishes a regularity and violation of this pattern by shortened SOA (early SOA deviant) or standard tone omission elicits MMN (Tse & Penney, 2007; Tse, Tien, & Penney, 2006). Alternating sound sequences, e.g., ABABAB, can also serve as the standard, and a repetition in the sequence, e.g., ABABB, elicits an MMN (Horvath, Czigler, Sussman, & Winkler, 2001). Abstract regularity established by ascending frequency tone pairs, using descending frequency tone pairs as the deviant, also elicits MMN (Saarinen, Paavilainen, Schoger, Tervaniemi, & Naatanen, 1992). For more complex rules involving the conjunction of two sound features (Paavilainen, Jiang, Lavikainen, & Naatanen, 1993), e.g., high pitch with high intensity, or low pitch with low intensity, violation of these rules, e.g., high pitch with low intensity, or low pitch with high intensity, also elicits an MMN. Infrequent stimuli may not elicit MMNs if the infrequent stimulus is part of a regular sequence. For example, in a cyclical repetition of the AAAAB sequence, the B does not elicit MMN. In contrast, when B is randomly inserted in a train of A, MMN is elicited, even when the probabilities of A and B are the same in the cyclical repetition condition and the random B condition (Sussman, Ritter, & Vaughan, 1998).

Although there are numerous evidence that provide support for the regularity violation model, understanding of the actual role of MMN in this framework is very limited. MMN is involved in violation detection, but no evidence shows that it is involved in maintaining/updating the predictive model. Moreover, there is no evidence that shows the presence of the feedback mechanism suggested by this model. Further, it is unclear whether any type of violation of regularity, even one involving complex context-dependent rules, can elicit an MMN.

Due to the distribution of the MMN scalp map, i.e. the mastoid inversion with the nose reference, it is predicted that there is a MMN generator on the superior surface of the temporal cortex. Lesion studies and brain imaging studies suggest a second MMN generator in the frontal cortex. Lesions to the temporal cortex (left or right or both) (Alain, Woods, & Knight, 1998) lead to a reduction in MMN, and larger MMN reduction has been found with right than left prefrontal lesion (Alain, Woods, & Knight, 1998; Alho, Woods, Algazi, Knight, & Näätänen, 1994).

In addition to the temporal generator, however, it is now clear that the MMN observed at the scalp typically includes activity from a frontal generator. Hints of the presence of this additional generator come from manipulations that influence activity at frontal electrodes to a greater extent than activity at central electrodes. Specifically, whereas factors modulating the temporal generator affect both the MMN measured in the frontal electrode and the associated inverted responses in the mastoid electrodes, some factors, such administration of alcohol (Jääskeläinen, Pekkonen, Hirvonen, Sillanaukee, & Naatanen, 1996), state of sleepiness (Sallinen & Lyytinen, 1997) and schizophrenia (Alain, Hargrave, & Woods, 1998) lead to a reduced MMN at frontal locations without affecting the mastoid inversion. This suggests that the frontal component is modulated independently of the temporal component. Equivalent current dipole (ECD) analysis showed that the scalp activity is best accounted for by a pair of bilateral dipoles

on the superior temporal plane pointing upwards and medially, and a right dipole in the frontal lobe (Giard, Perrin, Pernier, & Bouchet, 1990). fMRI studies using the passive oddball paradigm consistently found temporal cortex activity, but on occasion also activity in the frontal cortex supporting the presence of the frontal generator (Doeller, et al., 2003; Molholm, Martinez, Ritter, Javitt, & Foxe, 2005; Opitz, Rinne, Mecklinger, Cramon, & Schröger, 2002; Rinne, Degerman, & Alho, 2005). Numerous EROS studies have demonstrated temporal (Rinne, et al., 1999; Sable, et al., 2004; Sable, et al., 2007; Tse, Ng, Rinne, & Penney, 2009; Tse & Penney, 2007; Tse & Penney, 2008; Tse, Tien, & Penney, 2006) and frontal (Tse, Ng, Rinne, & Penney, 2009; Tse & Penney, 2007; Tse & Penney, 2008; Tse, Tien, & Penney, 2006) involvement in MMN. The details of the functional roles of the frontal and temporal generators will be discussed below.

**P300**

Another ERP component related to deviance detection is the P300, which refers to a positivity peaking at around 300-600ms (Donchin, 1981; Duncan-Johnson & Donchin, 1977). There are multiple subcomponents of the P300, i.e., P3a, P3b, parietally positive slow-wave, etc. (Fabiani, 2006; Pritchard, 1981). The classic P300 or the P3b has a centroparietal scalp distribution and is elicited by task relevant oddball stimuli. Its latency increases as the rare and frequent stimuli become more difficult to discriminate, indicating increased effort in evaluating the stimuli and increased processing time (Donchin & Coles, 1988; Duncan-Johnson & Donchin, 1982). The amplitude of P300 is sensitive to the probability of the stimuli, only if the stimuli are task relevant; P300 is absent when the stimuli are ignored (Duncan-Johnson & Donchin, 1977).

It was hypothesized that the P300 indicates context updating in working memory during the evaluation of stimuli where active responses are required (Donchin, 1981; Donchin & Coles,

1988; Duncan-Johnson & Donchin, 1982). In other words, the P300 is related to evaluating current stimuli and updating the current model or memory content for the processing of the next stimuli and future responses. The context updating hypothesis was supported by showing a relationship between the P300 amplitude and subsequent recall of the stimuli. If the P300 is related to working memory update, successful working memory update should correlate with better subsequent memory recall. Karis et al. (1984) investigated P3 and subsequent memory recall. Participants were divided into a rote mnemonic (rehearsal) and an elaborative mnemonic group based on the mnemonic strategies they used. Items that were subsequently recalled showed larger P300 than those that were not recalled in the rote mnemonic group, but this difference was not found in the elaborative mnemonic group. Amplitude of a slow positivity was correlated with subsequent recall in the elaborative mnemonic group. This result suggested that the rote mnemonic group repeatedly updated their memory (reflected by the P300), and items that were updated with more effort were better remembered. For the elaborative mnemonic group, because elaborative processing was more powerful than rehearsal, P300 amplitude did not predict subsequent recall. As predicted by the updating hypothesis, participants with low working memory capacity needed more effort (larger P300) to process changes in a sequence of stimuli than participants with high working memory capacity. This prediction was consistent with results observed in Brumback et al. (2005) that suggest participants with a low working memory span need more frequent updating.

Frontal P3, (also known as the P3a or the novelty P3), is elicited when novel stimuli are presented in an oddball sequence. It is also a positivity with a latency similar to that of the classic P3 but with a more frontal scalp distribution. Processing of the novel stimuli is different from that of the deviants in the oddball sequence, as there is no template or model for the unexpected

and rare novel stimuli. Fabiani and Friedman (1995) showed that the frontal P3 is elicited in the first deviant in each block of presentation, and the distribution shift to the posterior became the classic P3 when the deviant was repeated later in the sequence. This shift was observed in younger but not older adults, who instead appear to show frontal P3s both to novel and repeated rare stimuli. It was suggested that the frontal P3 is elicited when there is no memory template, and frontal lobe dysfunction occurring in aging may lead to problems in forming or maintaining the template.

The exact generators for the P3 have not been identified. It is believed that a distributed frontal-parietal network is involved (Halgren, Marinkovic, & Chauvel, 1998). Lesion studies showed that frontal lesions abolished the P3a while the P3b was not affected (Knight, 1984); and both the P3a and P3b amplitudes were reduced by temporal-parietal junction damage (Knight, Scabini, Woods, & Clayworth, 1989). Polich (2007) suggested that P3a is elicited when there is a shift in the focus of attention (reflecting activation of the frontal cortex), and P3 is elicited when the attention is used for updating memory (reflecting activation of the temporal-parietal junction). Some fMRI studies (Kirino, Belger, Goldman-Rakic, & McCarthy, 2000; Linden, 2005) could not separate the generators of P3a and P3b, but confirmed that the MFG and frontal-parietal network were activated in the "P3 process" in general. An EROS study also found frontal activities that corresponded to the time frame of the P3 effect (Low, Leaver, Kramer, Fabiani, & Gratton, 2006).However, the parietal region was not covered in the montage used in that study.

Several studies combined ERP and fMRI methods to localize the generators of the P300. By using fMRI to constrain the ERP dipole, Bledowski et al. (2004) found a frontal generator for the P3a and a temporal-parietal generator for the P3b. Calhoun et al. (2006) used independent

component analysis to combine ERP and fMRI, and showed initial activation in the auditory cortex and motor planning regions, followed by the auditory association cortex and motor execution regions. Later P3 response activated the temporal and medial frontal cortices. Similarly, Mantini et al. (2009) also identified activations of the motor, sensory and attention networks, during an active oddball paradigm, but only the ventral attention network, composed of the IFG and the parietal cortex, and the dorsal attention network, composed of the MFG/SFG and the parietal cortex, were related to the P300 amplitude. Benar et al. (2007) recorded EEG and fMRI simultaneously and used single trial ERP data in additional to the experimental conditions as the predictor for the BOLD effect. They found that single trial ERP predicted the fMRI response in the temporal-parietal junction and lateral frontal cortex. The frontal parietal network identified in these studies is consistent with the working memory network suggested by Fuster (Fuster, 2001; Fuster, 2006; Fuster, Bodner, & Kroger, 2000).

**N400 and P600**

Later ERP components have been shown to be sensitive to the deviance related to linguistic processes. Kutas and Hillyard (1980b) discovered a central parietal negativity at around 400ms, the N400,that is sensitive to semantic violations. The N400 is not specific to deviance detection but is elicited by stimuli carrying meaning, including not only written and oral language but also American Sign Language and pictures. Every word in a sentence elicits a N400 response; as the contextual constraint increases, the amplitude of the N400 decreases towards the end of the sentence. Incongruent words at the end of a sentence produce an increased N400 compared to congruent words at the end of a sentence, indicating that the N400 is sensitive to the violation of expectancy at the terminal word (Kutas & Hillyard, 1984). The N400 amplitude is also modulated by word frequency, position in a sentence, contextual constraints,

and cloze probability (Van Petten & Kutas, 1991). In other words, the amplitude of N400 is larger when a word deviates from the expectancy built up by the context or from experience, so that an increase in the N400 amplitude can be treated as a kind of regularity violation related to the semantic content. Increased N400 indicates semantic violation but not syntactic violation (Kutas & Hillyard, 1983).

Another ERP component, the P600, is sensitive to syntactic violation (Osterhout, 1997; Osterhout & Holcomb, 1992). P600 is a centroparietal positivity at 600-1000ms. It is elicited by syntactic violation (e.g. pronoun case violation, subject-verb agreement, and garden path sentences) and semantic violation (van Herten, Kolk, & Chwilla, 2005). The P600 is sensitive to both the probability and saliency of the syntactic violation. Some believe that the P600 is a P300 elicited in language processing paradigms (Coulson, King, & Kutas, 1998) . Generally, the P600 reflects processes in syntactic reanalysis and repairing (Hagoort & Brown, 2000).

Early lesion studies of patients with Broca's and Wernicke's aphasia suggested that the inferior frontal cortex is related to syntactic processing, whereas the left temporal cortex is involved in semantic processing. However, more recent ERP source analysis and fMRI work have shown a distributed network involving both the left inferior frontal and temporal cortices which is activated during language processing.

An MEG study (Halgren, et al., 2002) using the ECD method showed that the magnetic counterpart of the N400 was distributed in the frontal and temporal cortex. Sequential activation was found in the Wernicke's area at 250ms, the anterior temporal cortex at 270ms, the Broca's area at 300ms, and the anterior orbital and frontal pole at 370ms. The peak of the N400m involved the left anterior temporal, the perisylvian, and the dorsolateral prefrontal cortices.

Similar results were found by Simos et al. (1997). ERP studies with source localization suggested that the dipoles of the N400 were located along the perisylvian region including the Wernicke's area, superior temporal cortex, and temporal occipital parietal junction (D'Arcy, Connolly, Service, Hawco, & Houlihan, 2004). Frishkoff et al. (2004) showed the presence of early frontal dipoles and a later temporal dipole in N400. A combined ERP and fMRI study (Kiehl, Laurens, & Liddle, 2002) of N400 comparing incongruent and congruent sentence endings showed increased activation in the IFG, the MFG and the temporal cortex. These studies suggest the involvement of both frontal and temporal cortices in N400.

The anatomical basis for the frontal- temporal networks involved in language processing is provided by Parker et al. (2005). They found stronger connectivity of the ventral pathway which went from the Wernicke's to the Broca's area, connected through the uncinate fasciculus and medial STG in the language dominated hemisphere.

Slightly different frontal temporal networks are involved in semantic and syntactic anomalies. Ni et al. (2000), using fMRI, found that the network involved in semantic processing is more extensive than the one involved in syntactic processing. In addition, semantic anomalies activated the Wernicke's area more, while syntactic anomalies activated the Broca's area more. Newman et al. (2001) found that syntactic violation processing was associated with increased frontal activities, while semantic violation processing was associated with increased temporal and temporal parietal activations. Newman et al. (2003) further suggested that different parts of the frontal cortex were involved in semantic and syntactic processing. They found temporal cortex activation in both semantic and syntactic processing, but different parts of the inferior frontal cortex were activated in semantic processing (pars triangular) and syntactic processing (pars opercularis).

## Spatiotemporal Dynamics of Temporal and Frontal Cortices in Deviance Detection

In summary, despite different paradigms have been employed to study different kinds of deviance detection, activations in the frontal and temporal cortices are commonly found. However, the functional role of the frontal and temporal cortices in deviance detection remains unclear. Complicated cognitive processes can be made up of multiple stages and may involve feedback mechanisms. Identifying the brain regions involved in a cognitive process is the first step toward understanding its mechanisms. The next and very important step is to understand the spatiotemporal dynamics of the brain activities in a cognitive process. Event-related optical signal (EROS), a brain imaging method that has good spatial and temporal resolutions, is useful in this endeavor.

### Event-related optical signal (EROS)

It has been known for a long time that neuronal signals like the membrane potential and action potential are associated with changes in the optical properties of neurons (Cohen, 1973). However, only until recently has this property been exploited to image brain activity in humans non-invasively. Fast optical signal (FOS) and its event-locked counterpart, known as Event-related Optical Signal (EROS) (Gratton, Corballis, Cho, Fabiani, & Hood, 1995), use near infrared light to measure the optical changes associated with neuronal responses. The temporal resolution of EROS is in the millisecond range and its spatial resolution is in the centimeter range. It is based on measures of changes in photon "time of flight" from the source to the detectors, or intensity changes associated with changes in light scattering. When our brain is active, the associated cell swelling or conformational change in the cell membrane leads to a decrease in light scattering, thus light gets into a deeper layer of the grey matter before it is

scattered back to the detector. The light path is longer when the brain is active than when it is at rest. The increase in the length of the light path allows more light to be absorbed, leading to a decrease in light intensity. FOS/EROS is different from the more common optical imaging method, Near Infra-red Spectroscopy (NIRS), which measures slow hemodynamic signals with a temporal resolution of seconds, because NIRS measures variations in light intensity associated with oxy- and deoxy-hemoglobin concentration changes, for which light sources with two wavelengths are required.

Numerous studies have demonstrated the possibility of recording fast optical signals non-invasively in humans. These studies cover various fields in cognitive neuroscience and were published by several laboratories (Franceschini & Boas, 2004; Medvedev, Kainerstorfer, Borisov, Barbour, & VanMeter, 2008; Wolf, et al., 2002).Topics include vision (Gratton, Corballis, Cho, Fabiani, & Hood, 1995; Gratton, Fabiani, Elbert, & Rockstroh, 2003; Gratton, Sarno, Maclin, Corballis, & Fabiani, 2000), audition (Rinne, et al., 1999; Tse & Penney, 2007; Tse & Penney, 2008; Tse, Tien, & Penney, 2006), somatosensation (Franceschini & Boas, 2004; Maclin et al., 2005), and working memory(Low, Leaver, Kramer, Fabiani, & Gratton, 2006).

**Spatiotemporal dynamics and MMN**

In the following, MMN EROS studies are discussed to illustrate the importance of understanding the spatiotemporal dynamics in deviance detection. Specifically, how the spatiotemporal information is used to differentiate between three different hypotheses, deviance detection – attention switching, contrast enhancement, and response inhibition hypotheses of the MMN generators.

MMN was first suggested to be involved in pre-attentive deviance detection and possibly attention switching (deviance detection – attention switching hypothesis) (Näätänen & Michie, 1979). Two sets of generators were proposed: the bilateral temporal generators that are associated with the establishment of memory trace and deviance detection, and the right frontal generator that is associated with attention triggering or attention switching after deviance detection (Näätänen & Michie, 1979). The deviance detection – attention switching hypothesis predicted a specific spatiotemporal pattern of brain activity: if the frontal source is responsible for attention switching and the temporal source is responsible for deviance detection, temporal followed by frontal activation would be expected.

An ERP/MEG source localization MMN study (Rinne, Alho, Ilmoniemi, Virtanen, & Näätänen, 2000) attempted to establish this spatiotemporal relationship. In ERP, maximum cortical currents for sound duration decrement were found in the temporal and inferior frontal area. More interestingly, the temporal source leads the frontal source by 8 ms. However, due to spatial smearing, the frontal and temporal generators were not fully separated (Rinne, Alho, Ilmoniemi, Virtanen, & Näätänen, 2000) and this 8-ms temporal generator lead may be an underestimation.

The deviance detection – attention switching hypothesis is challenged by fMRI studies (Opitz, Rinne, Mecklinger, Cramon, & Schröger, 2002; Rinne, Degerman, & Alho, 2005) using a parametric design. Deviant stimuli with increasing degrees of deviance are used. As predicted by the deviance detection – attention switching hypothesis, a larger deviance is more salient and likely to initiate attention switching. If the frontal and temporal cortices are responsible for attention switching and deviance detection respectively, activities in both cortices should increase with the degree of deviance. However, Opitz et al. (2002) only showed a linear increase

in the temporal cortex activity. In the frontal cortex, an inverted quadratic trend of brain response was associated with increasing degrees of deviance. An alternative model, contrast enhancement (Opitz, Rinne, Mecklinger, Cramon, & Schröger, 2002), was then proposed. This model suggested that when the deviant is ambiguous and difficult to distinguish from the standard, the frontal generator is activated to amplify the difference between the deviant and standard. As the medium deviant was the most ambiguous, the largest frontal activity was observed.

Similar to Opitz et al. (2002), Rinne et al. (2005) applied a parametric design, but used deviants obtained with duration decrements to reduce the N1's influence on MMN. The result showed a linear trend of brain activity in the temporal cortex but an inverse linear trend of brain activity in the frontal cortex. Again, this result is incompatible with the deviance detection-attention switching hypothesis. Rinne et al. (2005) suggested a third hypothesis, response inhibition, about the functional role of the frontal cortex in pre-attentive change detection. This model suggested that the frontal cortex is responsible for implementing inhibition on down-stream attention or response mechanisms. Thus, an inverse linear trend in the frontal cortex is expected, as the smaller the deviant, the less likely that it requires further responses.

Both the contrast enhancement and response inhibition hypotheses can explain the modulation of brain responses by the degree of deviance. However, limitations of these hypotheses are revealed once the spatiotemporal dynamics of the frontal and temporal cortices are shown through EROS.

Tse and Penney (2008) adopted a parametric design with frequency deviants similar to those used in Opitz et al. (2002). A linear trend was found in the temporal activity, while an inverse quadratic trend was found in the frontal activity. Although the temporal and frontal

modulation by the degree of deviance was similar to that of Opitz et al. (2002), the temporal dynamics revealed by EROS cast doubts on the contrast enhancement hypothesis: temporal followed by frontal activities were found for the small and the medium deviant conditions. As suggested in Tse and Penney (2007), if the frontal activation is related to contrast enhancement while the temporal activation is related to deviance detection, contrast enhancement should take place before deviance detection, i.e. frontal activity should occur before the temporal activity. However, temporal followed by frontal activity will be observed in the attention switching (Näätänen & Michie, 1979) or response inhibition hypothesis (Rinne, Degerman, & Alho, 2005).

Tse and Penney (2007) and Tse et al. (2009) did in fact observe both "frontal followed by temporal" and "temporal followed by frontal" patterns, but only with stimuli that deviate from the standard in the temporal dimension. Tse and Penney (2007) found a frontal-temporal-frontal activation pattern by deviants with an early SOA, but a temporal followed by frontal activation pattern by omission deviants. It was argued that the omission deviants were more salient than the early SOA deviants, thus no contrast enhancement was required for omission deviants. On the other hand, early SOA deviants are ambiguous stimuli that require contrast enhancement. Indeed, early frontal cortex activity was found in this condition. This study suggests that the frontal cortex may have different functional roles across time. The early frontal activity is related to contrast enhancement, which leads to temporal activity for deviance detection, whereas later frontal activity may be related to inappropriate representation or response inhibition as suggested by Rinne et al. (2005).

With the relatively good resolution in both spatial and temporal dimensions provided by EROS, it is possible to statistically test how the frontal temporal activation pattern is modulated by the degree of deviance. Tse et al. (2009) applied a lagged correlation path model , i.e., a type

of dynamic structural equation model (SEM), to study the correlation of temporal and frontal activations across time (see Rykhlevskaia, Gratton, & Fabiani, 2008). An MMN paradigm with duration decrement deviants was used, as in (Rinne, Degerman, & Alho, 2005). The results showed temporal-frontal, frontal-temporal-frontal, and temporal-frontal activation sequences in large, medium, and small deviant conditions, respectively. The lagged correlation path modeling suggested that activation of the frontal temporal network was dynamically modulated depending on the degree of deviance. In addition, further analysis on the BOLD equivalent slow optical signal showed similar results as Rinne et al. (2005).

The EROS studies reviewed above illustrate the importance of uncovering the spatiotemporal dynamics of the frontal and temporal cortices in pre-attentive change detection. Specifically, spatiotemporal information can be used to guide the development of a theory of how a cognitive process is implemented in the brain.

### Current Thesis

In the current thesis, three experiments are proposed to investigate the involvement of the frontal temporal network in deviance detection. Although EROS has been used to investigate a wide range of cognitive processes, especially deviance detection, it has not been used to study semantic and syntactic violations during sentence comprehension. The first study aimed at investigating the dynamics of the frontal temporal network in detecting deviance during sentence processing. A lot of effort has been made on how deviance is detected. However, the establishment of regularity is equally important. In the second study, representations of contingency rules governing a sequence of stimuli were studied. Although MMN is one of the best researched areas with EROS, all of the MMN EROS studies to date focused on detecting

physical changes of the stimuli in the auditory domain. The last study investigated how the frontal temporal network is involved in detecting deviance elicited by an auditory illusion. Specifically, audiovisual MMN was investigated with McGurk deviants.

# Chapter 2

## Imaging Cortical Dynamics of Language Processing with the Event-related Optical Signal (EROS)

### Introduction

Language processing involves the rapid interaction of multiple brain regions. The study of its neurophysiological bases would therefore benefit from neuroimaging techniques combining both good spatial and good temporal resolution. Here we use the event-related optical signal (EROS), a recently developed imaging method, to reveal rapid interactions between left superior/middle temporal cortices (S/MTC) and inferior frontal cortices (IFC) during the processing of semantically or syntactically anomalous sentences. Participants were presented with sentences of these types intermixed with nonanomalous control sentences and were required to judge their acceptability. ERPs were recorded simultaneously with EROS and showed the typical activities that are elicited when processing anomalous stimuli: the N400 and the P600 for semantic and syntactic anomalies, respectively. The EROS response to semantically anomalous words showed increased activity in the S/MTC (corresponding in time with the N400), followed by IFC activity. Syntactically anomalous words evoked a similar sequence, with a temporal-lobe EROS response (corresponding in time with the P600), followed by frontal activity. However, the S/MTC activity corresponding to a semantic anomaly was more ventral than that corresponding to a syntactic anomaly. These data suggest that activation related to anomaly processing in sentences proceeds from temporal to frontal brain regions for both semantic and syntactic anomalies. This first EROS study investigating language processing shows that EROS can be used to image rapid interactions across cortical areas.

Language processing evolves rapidly over time and involves multiple brain regions. Lesion studies have long identified both superior/middle temporal cortices (S/MTC) and inferior frontal cortices (IFC) in the left hemisphere (often referred to as Broca's and Wernicke's areas) as critical regions for language processing (Goodglass, 1993). However, the way in which these areas interact is subject to debate (Bushell, 1996; Caramazza, Capitani, Rey, & Berndt, 2001; Kaan & Swaab, 2002). Neuroimaging methods may provide information about the dynamics of these interactions.

Because most brain-imaging techniques have either good temporal or good spatial resolution, but not both, different aspects of language processing have been investigated with different methods. Because of their exquisite temporal resolution, event-related brain potentials (ERPs) have been the method of choice for studying the temporal aspects of language-related processes. The N400 and P600 components of the ERP have been found to be related to the processing of semantic (Kutas & Hillyard, 1980a)and syntactic (Osterhout & Holcomb, 1992) anomalies, respectively. However, because of the relatively low spatial resolution of ERPs, it is difficult to link the electrophysiological responses to brain regions, which is particularly problematic for language because the relevant areas may be in close proximity.

In contrast, functional magnetic resonance imaging (fMRI) provides more precise spatial information, and studies using this method also have implicated both the left IFC and S/MTC in both semantic and syntactic processing (Kaan & Swaab, 2002). A recent review of such studies by Hagoort (2005) led to the proposal that the lexical properties of each perceived word are retrieved in S/MTC and then sent to the IFC, which integrates information across multiple words over time.

Hagoort's model emphasizes a functional specialization within frontal regions, with more posterior and dorsal areas related to the integration of lexical-syntactic information and more anterior and ventral areas related to lexical-semantic integration. Hagoort further speculated that a similar pattern might extend to temporal regions where the syntactic and semantic properties of words are initially retrieved, although the evidence so far is sparse. However, fMRI alone does not have sufficient temporal resolution to determine the relative timing of different activities occurring during language processing, such as the flow of information from temporal to frontal regions, which typically unfolds over just fractions of seconds.

Analysis of the neural basis of sentence processing requires separating, in both space and time, the activity observed in each of the relevant regions. Several methods have been used to address this need, including magnetoencephalography (MEG) (Halgren, et al., 2002) and intracranial recordings (McCarthy, Nobre, Bentin, & Spencer, 1995). In the current study, we used a recently developed functional brain-imaging method, the event-related optical signal (EROS) (Gratton & Fabiani, 2007). EROS identifies changes in the light scattering properties of cortical tissue related to neuronal activity (i.e., fast optical signals, as opposed to optical correlates of hemodynamic signals) (Rector, Carter, Volegov, & George, 2005) by using near infrared light. With high spatial and temporal sampling, it is possible to obtain EROS images with spatial and temporal resolutions on the order of a few millimeters and milliseconds, respectively.

EROS has been used to study basic sensory processes in the visual (Gratton, Corballis, Cho, Fabiani, & Hood, 1995), auditory (Rinne, Alho, Ilmoniemi, Virtanen, & Näätänen, 2000), and somatosensory (Maclin, Low, Sable, Fabiani, & Gratton, 2004) modalities, as well as higher cognitive functions, such as preattentive change detection (Rinne, Alho, Ilmoniemi, Virtanen, &

Näätänen, 2000; Tse & Penney, 2007; Tse, Tien, & Penney, 2006), attention modulation, and target detection (Low, Leaver, Kramer, Fabiani, & Gratton, 2006). However, no EROS study of language processing has been conducted thus far. In this study, we show how EROS can identify two adjacent, but distinct, areas in S/MTC that are involved in the processing of semantic and syntactic anomalies and are associated with scalp-recorded ERP activity (N400 and P600, respectively). We also study the relative timing of frontal and temporal EROS activity associated with these anomalies.

**Methods**

Sixteen right-handed native English speakers (11 females, ages18–30) participated in this study after providing informed consent. Procedures were approved by the campus Institutional Review Board. Participants were presented with sentences of five to eight words, word by word at the center of a computer screen, and instructed to make a yes/no decision about whether each sentence was both well formed and sensible (henceforth called ‘‘acceptable’’) after its completion (see Figure 2.1). Each participant saw 864 sentences (528 acceptable, 336 unacceptable). Among the 336 unacceptable sentences, 144 became semantically anomalous at the final (critical) word (e.g., ‘‘The hungry child ate the floor’’), 96 exhibited grammatical violations of subject–verb agreement (e.g., ‘‘If work isn’t done, it pile. . . ’’), and 48 had grammatically incorrect pronoun case (e.g., ‘‘My mother promised to buy I. . . ’’). For the ungrammatical sentences, the critical words were at positions 4–8. To prevent expectations that semantic anomalies could only occur in the final position, 48 filler sentences contained semantically anomalous words in positions 3 to 7. Of the 528 acceptable sentences, 288 were the matched controls for the semantic (144), syntactic subject–verb agreement (96), and pronoun

case (48) anomaly conditions. The remaining 240 were included so that participants would expect that most sentences would be correct, which they generally are in normal comprehension.

In the major analyses presented here, ungrammatical sentences with subject–verb and pronoun–case disagreement were combined to maximize statistical power, and only those semantically anomalous sentences whose critical word was the final one were included because they were the only semantic anomalies with matched control versions. A secondary analysis was conducted to examine midsentence semantic anomalies. There were two complementary lists of stimuli (i.e., acceptable sentences in list 1 were the correct versions of unacceptable sentences in list 2, and vice versa), and each participant saw only one list. The mean lengths and frequencies of the critical words in acceptable and unacceptable sentences did not differ. EROS and ERPs were recorded simultaneously. The semantic and syntactic conditions described throughout this paper refer to the brain activity elicited by critical words in the acceptable sentences subtracted from that for the corresponding words in the unacceptable sentences, time-locked to the eliciting word onset.

**ERP Recordings**

The EEG was recorded with gold electrodes at four scalp locations based on the 10/20 system (Fz, Cz, Pz, and right mastoid) referenced to the left mastoid (with an average mastoid reference computed offline). Four electrodes, one above and one below the right eye and two at the outer canthi of each eye, were used for vertical and horizontal EOG recording. Electrode impedance was 5 kOhms. The EEG was filtered online by using a 0.01- to 30-Hz bandpass sampled at 100 Hz, filtered offline with a 0.1- to 20-Hz bandpass, segmented by using 1500-ms epochs time-locked to the onsets of critical words (with 200-ms prestimulus baselines), and

averaged according to the stimulus condition. Ocular artifacts were corrected (Gratton, Coles, & Donchin, 1983), and epochs containing other EEG artifacts were removed from the analysis.

**EROS Recordings**

EROS data were recorded using a frequencydomain oximeter (Imagent; ISS, Inc., Champaign, IL). Nearinfrared light (830 nm) from laser diodes modulated at 220 MHz was conducted to the participant's head by optical fibers (sources). Light that scattered through the head was carried by detector optical fibers to photo-multiplier tubes used for the measurements. The photo-multiplier tubes were modulated at 220.00625 MHz generating a 6,250-Hz heterodyning frequency (i.e., crosscorrelation frequency). The output current was Fast Fourier transformed, and relative phase-delay measures (in degrees) were computed.

The light source and detector fibers were held in place by using a custom-built head-mount system (Tse & Penney, 2007; Tse, Tien, & Penney, 2006). Two montages (i.e., two sets of light source/detector configurations) (Figure 2.2) were used to interrogate the left S/MTC and IFC. The configurations of the two montages were identical, but one of them was positioned 1.7cm anterior to the other. This arrangement allowed higher spatial sampling compared to using only a single montage. Each montage comprised 128 source-detector pairs (8 detectors, each multiplexed with 16 of 24 light sources). The sampling frequency was 39.0625 Hz (i.e., it took 25.6 ms to cycle through the 16 multiplexed channels). Montage order was counterbalanced across subjects.

For coregistration with individual subjects' anatomy, the locations of sources and detectors, the nasion and preauricular points, and another 185 random points over the face and scalp of each person were digitized with a Fastrak 3Space 3D digitizer (Colchester, VT). A T1-

weighted brain anatomical (MPRAGE) MRI was obtained for each participant by using a 3T Siemens Magnetom Allegra Headscanner (New York, NY). The nasion and preauricular points were marked by vitamin E pills in the MRI scan for coregistration of the functional optical data with the structural MRI. The coregistration was based on a successive application of fiducial and scalp-fitting methods that have been shown to reduce the coregistration error to less than 5 mm (Whalen, Maclin, Fabiani, & Gratton, 2008). The individually coregistered data were then Talairach-transformed (Talairach & Tournoux, 1988) before statistical analyses.

Optical data were preprocessed by correcting for phase wrapping, normalizing, pulse correcting (Gratton & Corballis, 1995), filtering with a 0.01- to 8-Hz bandpass filter, and then averaging for each channel, time point, condition, and subject by using a 200-ms prestimulus baseline. Channels with phase standard deviations greater than 210 ps and/or a source detector distance outside the 15- to 75-mm range (about16.5%) were excluded from the analysis. The same criteria were applied to all conditions.

Special-purpose software (Opt-3D) (Gratton, 2000) was used to analyze the averaged data. With the assumption that the light path in the brain is similar to that obtained in a homogenous medium, the optical signal for a given voxel was calculated as the mean value of the channels overlapping at that particular voxel (Wolf, et al., 2000). The montage was designed to generate a high degree of overlap between the volumes interrogated by different channels (greater than 50% overlap between adjacent channels). This approach was useful for the spatial reconstruction process, resulting in an improved signal-to-noise ratio and spatial resolution, allowing us to examine focal EROS responses. T statistics for the phase data were calculated at the group level for each voxel and then converted to Z scores. Statistical maps of the optical signal for each data point were generated by 3D reconstruction of the Z scores on a template

brain in Talairach space, with an 8-mm spatial filter according to the location information from the coregistration procedure. The ROIs for the statistical analysis of brain activation in S/MTC and IFC were based on previous brain-imaging studies (Caplan, Alpert, & Waters, 1998; Kiehl, Laurens, & Liddle, 2002; Kuperberg, et al., 2000; Meyer, Friederici, & von Cramon, 2000; Moro, et al., 2001; Newman, Pancheva, Ozawa, Neville, & Ullman, 2001). The S/MTC and IFC ROIs were tested independently in the analyses, and only the statistically significant peak activities are reported here (more details regarding EROS recording and analysis methods can be found in Gratton and Fabiani, 2007 and Gratton, et al., 2006).

## Results

### Semantic vs syntactic anomalies

Subjects classified most sentences correctly: semantically acceptable (94%), semantically unacceptable (95%), syntactically acceptable (88%), and syntactically unacceptable (86%). Only correct trials were included in the following analyses.

To test for the presence of N400 and P600 effects in the semantic and syntactic conditions, mean ERP amplitudes were computed on the difference waveforms in time windows from 200 to 500 ms (N400) and from 500 to 1500 ms (P600). The N400 and P600 time windows were selected based on previous ERP language studies (Kutas & Federmeier, 2000; Osterhout, 1997; Osterhout & Holcomb, 1992). The mean values were tested against baseline by using one-sample t tests (one-tailed) on the Pz electrode. A significant increase in N400 (Figure 2.3, Upper) (mean = -6.06 μV, SD = 2.46, $t(15) = -9.88$, $p < .001$, peaking at 420 ms) was found in the semantic condition, and a significant increase in P600 (Figure 2.3, Lower), peaking at 860 ms, was found for the syntactic condition (mean = 5.58 μV, SD = 3.07, $t(15) = 7.26$, $p < .001$). There

were no differences in either the P600 time window in the semantic conditions (mean = .73 μV, SD = 2.73, $t(15) = 1.07$, *ns*) or the N400 time window in the syntactic conditions (mean = .36 μV, SD = 1.40, $t(15) = 1.02$, *ns*).

Statistical maps of EROS showed significant effects (i.e., increased phase delay for anomalous critical words) in left S/MTC and IFC for both the semantic and syntactic conditions. Table 2.1 summarizes the latency[1], Talairach coordinates, peak Z scores, critical Z (with $p < .05$ corrected for multiple comparisons) (Poline, Worsley, Evans, & Friston, 1997), corresponding brain region, and Brodmann's area (BA) of the largest statistically significant peak EROS activity in each time interval. Activity at similar locations and adjacent time intervals was regarded as belonging to the same temporospatial cluster of activity. Figure 2.4 a and b shows the most significant peak EROS response of each cluster for the semantic and syntactic conditions, respectively. The relative position of each peak response is shown on a left lateral view of the brain in Figure 2.4c (Upper, semantic condition; Lower, syntactic condition).

As shown in Table 2.1 and Figure 2.4, a pattern with S/MTC activation followed by IFC activation was elicited in both the syntactic and semantic conditions, beginning at 179 ms in the semantic condition and at 256 ms in the syntactic condition. In the semantic condition, this activation pattern recurred multiple times, suggesting oscillatory activity. This oscillatory activity was different from the rhythmic activities, e.g., alpha or gamma oscillations in EEG, and instead suggested re-entry or feedback/feedforward of information between the frontal and temporal cortices. Note that the two regions of interest (ROIs) were analyzed independently, and

[1] The latencies given in Table 2.1 refer to the beginning of each 25.6 ms sampling interval.

that the statistically significant peak for each individual ROI at each time point was selected, rather than the highest value across both ROIs.

Figure 2.5 presents a direct comparison of the EROS responses in the semantic (small circles) and syntactic (small squares) conditions. This figure indicates that different areas in the temporal and frontal lobes were involved in the semantic and syntactic conditions up to about 665 ms, whereas similar areas were involved in the semantic and syntactic conditions after 665 ms. A more ventral anterior/middle temporal (dark ellipse) region was involved in early semantic processing, whereas a more dorsal posterior temporal region (light ellipse) was involved in early processing in the syntactic condition. A more extensive area of the frontal cortex was involved in the semantic (dark square) than in the syntactic (light square) condition across time[2]. In addition, two regions were activated in both conditions at similar times: a dorsal temporoparietal area (BA 40/41) activated between 665 and 844 ms (yellow rectangle) and an inferior/middle frontal area (BA46) activated about 927 ms (red rectangle). These data suggest that, early on, distinct areas respond to semantic and syntactic anomalies, but that at longer latencies some regions are involved in responding to both anomaly types. A separate analysis conducted on midsentence semantic anomalies (see analysis below) showed a similar, albeit weaker, temporospatial pattern of activity elicited by semantic violations, suggesting that these differences are not due to word position in the sentence.

[2]Alternatively, the IFG regions activated in the semantic and syntactic conditions may be of similar extent. However, because the semantic activity is more extended in time, it is more likely for different points to become active at different times.

To specifically investigate the relationship between EROS and ERP signals, stepwise multiple regressions[3]§ were conducted to determine which of the observed optical effects, if any, predicted the amplitudes of the N400 and P600 ERP components across subjects. EROS activity elicited in the semantic condition in both S/MTC (at 179 ms, $\beta = 0.76$, $p < .001$ and 384 ms, $\beta = -2.27$, $p < .005$) and IFC (at 512 ms, $\beta = 1.37$, $p < .05$) predicted the N400 ERP effect [$R^2 = 0.884$, $F(3,7)= 17.88$, $p < .001$] (Table 2.2). The P600 effect was predicted by the EROS in S/MTC (at 819 ms, $\beta = 0.46$, $p < .005$ and 914 ms, $\beta = -0.52$, $p < .05$) elicited in the syntactic condition [R2 $R2 = 0.608$, $F(2,12)= 9.31$, $p < .005$] (Table 2.2). No statistically significant relations were found by using optical signals from the syntactic conditions to predict the N400 effect or from the semantic condition to predict the P600 effect (Table 2.2). Taken together, these analyses suggest a double dissociation in the EROS–ERP prediction pattern. Only EROS signals elicited in the semantic condition predict the N400 effect, and only EROS signals elicited in the syntactic condition predict the P600 effect.

Among all significant predictors, the EROS activity elicited in the semantic condition in S/MTC at 384 ms was the only predictor that increased with larger N400s (the sign for this predictor is negative because the N400 becomes more negative as its amplitude increases), whereas the EROS activity elicited in the syntactic condition in S/MTC at 819 ms was the only predictor that increased with larger P600s. Thus, once the effects of other significant predictors were removed, the S/MTC EROS activity at 384 ms increased as the size of the N400 effect

[3]The statically significant optical effects reported in Table 2.1 were entered into the model as candidate predictors. Because EROS responses extended across time, responses at adjacent time points are highly correlated with both each other and the ERP effect. If all EROS responses were entered simultaneously into the regression model, there would be a high degree of multicollinearity resulting in unstable parameter estimation. This result was avoided by using a stepwise procedure, in which only the sets of EROS effects that best correlated with the ERP effects were selected. The results were reported in the final models (Table 2.2). The predictors selected into the final model by the stepwise procedure encompassed the duration of the N400 and P600 effects.

(peaking 36 ms later at 420 ms) increased, and the S/MTC EROS activity at 819 ms increased as the size of the P600 effect (peaking 41 ms later at 860 ms) increased. Interestingly, the other significant predictors, which were not as close in time to the N400 or P600 peaks, were opposite in sign from the ones just described (i.e., positive for the N400 and negative for the P600). Thus, when the predictors increased, N400 or P600 amplitude decreased, suggesting that they may reflect neural activities modulating the N400 or P600 effects.

To determine whether the location of the optical activity in S/MTC differed between semantic and syntactic conditions, the location of the peak EROS activity corresponding to the N400 effect (semantic S/MTC at 384 ms) was compared with the location of the peak EROS activity corresponding to the P600 effect (syntactic S/MTC at 819 ms) (Fig. 6 a and b) by using a jackknife procedure and multivariate t test. The peak locations for the two conditions differed (Hotelling's $T^2$=11.02, $p$<.05). The peak of the S/MTC optical signal obtained in the semantic condition was 17 mm inferior to that obtained in the syntactic condition [$t$(15)=3.21, $p$<.005], whereas there was no difference along the anterior–posterior dimension [$t$(15)=0.59, *ns*]. Thus, the combined temporal and spatial resolutions of EROS allowed us to separate the optical signals corresponding to the N400 and P600 in the S/MTC region, although they were _2 cm apart. The relative location of these signals in the semantic and syntactic conditions is consistent with Hagoort's (2005) speculation that regions handling lexical syntactic information may be more dorsal than those handling lexical-semantic information in S/MTC. At latencies exceeding 665 ms, some common areas were activated by syntactic and semantic anomalies. Within both the superior temporal gyrus (latency, 844 ms) and inferior frontal gyrus (IFG) (latency, 972 ms), there were no reliable differences between the locations of activity for the syntactic and semantic

contrasts. Some of this late activity is probably related to the behavioral response, which is the same for the two kinds of anomalies.

### Midsentence anomalies

A question addressed in the current study is whether the brain activation pattern in response to semantic anomalies differed from the response to syntactic anomalies. Because the semantic anomalies were in sentence-final position, whereas the syntactic anomalies were sentence-medial, it is important to demonstrate that similar regions are activated at similar latencies during the processing of sentence-medial semantic anomalies as during sentence-final ones. Because we included a set of filler sentences containing sentence-medial semantic anomalies in the same range of positions as the syntactic anomalies, we conducted supplementary analyses.

We originally chose to manipulate and analyze semantic deviance only at the end of the sentence because it has been shown that this position typically maximizes N400 effects, especially in sentences as short as the ones we used here, and we wanted to have the greatest possible power for this initial EROS study. However, we included a set of 48 filler sentences that had semantically deviant words in positions that were the same as those of the syntactic deviants, albeit without exactly matching controls. They were included to prevent subjects from building expectations that the semantic deviant could only occur at the end of the sentence. By analyzing these sentences, we were able to compare the optical and ERP effects for these words to those from the final words in the sentences.

It is important to note that word length and frequency are not as well matched in this comparison as in the other comparisons. The critical words in the control versions of the

sentences with sentence-final semantic anomalies were actually perfectly matched overall in length and frequency with the sentence-final semantically anomalous words, because the sentence-final semantic anomalies were created by simply re-pairing sentence frames and final words. In the syntactic anomalies, the critical word in the good and bad versions varied only by whether the verb was singular or plural or by the case of the pronoun (e.g., I vs. me). There were equal numbers of singular and plural syntactically anomalous verbs, and the same goes for pronouns with the particular case variations we used. Thus, within anomaly type, word length and frequency of the critical and control words were well matched. In contrast, in the new analysis of sentence-medial semantic anomalies reported here, we used the control words from sentence-medial positions in the syntactic anomalies as the control condition. These were verbs and pronouns, whereas the midsentence semantic anomalies were mostly nouns (plus a couple of verbs or words of other types). However, although these words were not intentionally matched, word length and frequency were still fairly close overall (average word length = 4.54 vs. 5.25 letters, and average word frequency = 2.27 vs. 1.85 log frequency). One benefit of using this control is that exactly the same baseline was subtracted from both semantic and syntactic anomalies, the two conditions we intend to compare to each other. Thus, we believe this is an appropriate comparison to make, and if this subtraction yields results that are consistent with those of the sentence-final semantic anomalies, we would argue that the effects are not due just to word position.

The brain activation patterns elicited by semantic anomaly occurring in the middle and at the end of the sentence are compared in Figure 2.7. There is substantial similarity in the temporospatial brain activation patterns elicited by semantic anomaly in the two sentence positions. The same S/MTC or IFC activities, with similar latencies, were shown in both

sentence-final and sentence-medial conditions (green dotted circles). However, the effects observed for the sentence-medial positions were considerably weaker than those observed for the sentence-final position (note that the scale used in this figure is smaller than that in Figure 2.4). The biggest difference between sentence positions seems to be in the late frontal activity at 972 msec, which is probably related to the fact that participants have to respond soon after the sentence-final anomalies appear, whereas they do not respond until several words after the sentence-medial anomalies.

The similar but weaker sentence-medial EROS effect is consistent with ERP studies where the N400 effect for semantic anomaly is typically bigger toward the end of the sentence. As contextual constraint builds up across a sentence, the final word in acceptable sentence versions becomes fairly predictable and thus easier to integrate with the rest of the sentence, leading to smaller N400s in the control condition. The same accrual of contextual constraint leads to a larger N400 when the final word is anomalous because the more context, the worse the fit of the anomalous word. This combination tends to yield large differences in N400 amplitude between sentence-final anomalous and control conditions, and that was true in our study. We analyzed the ERP N400 effect for the sentence-medial semantic anomalies (see Figure 2.8.) and found a smaller but significant semantic anomaly effect in the middle of the sentence (sentence-final position, mean = -6.67 mV, SD = 2.58, $t(15) = -10.335$, $p < .001$; sentence-medial position, mean = -1.83 mV, SD = 2.81, $t(15) = -2.599$, $p < .05$). This analysis used a 350- to 500-ms time window, which was slightly smaller than the 300- to 500-ms time window used in the analyses reported in our paper because the sentence-medial N400 was narrower).

The similarity of the spatiotemporal EROS pattern for sentence-medial and sentence-final anomalies suggests that processing of semantic anomaly is associated with activity in more

ventral temporal regions than the processing of syntactic anomaly regardless of sentence position.

**Discussion**

EROS revealed the dynamic activation of partially overlapping distributed networks associated with the processing of semantic and syntactic anomalies. The location of these temporal-frontal networks is consistent with previous fMRI studies (Bookheimer, 2002; Bookheimer, 1996; Friederici, 2002; Friederici, Ruschemeyer, Hahne, & Fiebach, 2003; Hagoort, Hald, Bastiaansen, & Petersson, 2004; Just, Carpenter, Keller, Eddy, & Thulborn, 1996; Ruschemeyer, Fiebach, Kempe, & Friederici, 2005), which suggest that frontal regions (BA44, BA45, and BA47), temporal regions (BA37, BA38, BA40, BA41, BA21, and BA22), and the precentral gyrus (BA6) are involved in language processing. This posterior–anterior flow of activity for semantic anomaly processing also has been observed by using MEG (Halgren, et al., 2002).

We interpret these activation patterns in terms of a model proposed by Hagoort (2005), in which comprehension involves the retrieval of lexical information in temporal regions and the subsequent integration of that information across words in frontal regions. In addition, we bring in a proposal by Federmeier (2007) that left-hemisphere language areas work in a top–down predictive mode. Specifically, we suggest that the integration process in IFC includes the generation of predictions about upcoming words, which are communicated to temporal memory areas. These predictions concern both semantic features (e.g., that the upcoming word will designate something edible) and syntactic features (e.g., that it will be a plural verb). We speculate that it is the failure of the input to match these predictions that produces the earliest

EROS effects in temporal regions (200–400 ms) in different S/MTC locations for semantic and syntactic information. Further, our syntactic anomalies, being simple number or case mismatches, may be easier to integrate in frontal regions because they require correcting only a single feature, so there is little differential frontal response to the syntactic anomalies, aside from the late frontal activation we ascribed to the behavioral response (972 ms). In contrast, the attempt to integrate the semantic anomalies leads to interpretations that cannot be easily corrected (e.g., children eating floors), and hence there is an anomaly-sensitive response in IFC (e.g., 512 ms), as well as a need to further consult lexical memory in temporal cortex (e.g., S/MTC activation at 665 ms) possibly to check for accuracy of the retrieval process. The result is oscillation between the anterior and posterior language areas in response to semantic anomaly.

Our data suggest spatial segregation of lexical-semantic and lexical-syntactic processing in the temporal lobes, just as previous fMRI studies (Hagoort, Hald, Bastiaansen, & Petersson, 2004; Petersson, Forkstam, & Ingvar, 2004; Saur, et al., 2008) have found spatial segregation of semantic and syntactic processing in the frontal lobes. Interestingly, the pattern in temporal regions, with the response to syntactic anomaly more posterior and dorsal than that to syntactic anomaly, is consistent with Hagoort's (2005) speculation that the pattern observed in frontal regions may extend to temporal regions as well.

With regard to frontal cortex, the same region of IFG is activated between 972 and 998 ms by the two anomaly types, possibly reflecting participants' ultimate judgment that the sentence is incorrect in some way. However, a similar (although slightly more inferior) (Hagoort, Hald, Bastiaansen, & Petersson, 2004; Petersson, Forkstam, & Ingvar, 2004; Saur, et al., 2008) region is already activated at a latency of 563 ms in the semantic condition, suggesting that the frontal cortex responds to a semantic anomaly earlier than it does to syntactic anomaly, at least

for the kinds of sentences we used. The current study also suggests that the more extensive involvement of frontal regions in the semantic conditions observed in fMRI studies (Bookheimer, 2002) may reflect the superimposition of adjacent frontal activations across time.

In addition to IFG, however, the data show other regions involved in the processing of both syntactic and semantic anomaly: Dorsal superior temporal gyrus, which is proposed to be involved in syntactic processing, also is involved in processing semantic anomalies at 665 and 844 ms poststimulus (i.e., within the range of the ERP P600 effect). The hypothesis outlined earlier about IFC sending predictions about semantic and syntactic features of upcoming words to S/MTC is one possible explanation for this region's involvement in both the semantic and syntactic conditions.

In this study, we established a correspondence between the N400 and P600 ERP effects and EROS activity in S/MTC. Previous EROS studies (Low, Leaver, Kramer, Fabiani, & Gratton, 2006; Rinne, Alho, Ilmoniemi, Virtanen, & Näätänen, 2000; Tse, Tien, & Penney, 2006) have shown temporal correspondence between optical and ERP signals, but in the current study we also showed that EROS amplitude is correlated with ERP amplitude. The optical data accounted for a high proportion of variance in the ERP components (88% for the N400, 61% for the P600). Further, the ERP–EROS correspondence was specific to each condition, as revealed by the double dissociation in the prediction pattern. This finding rules out accidental factors, such as differences across subjects in skull thickness, which may lead to a spurious relationship between ERP and EROS, and points instead to stimulus-specific neuronal factors. However, we are not attempting to locate the generators of N400 and P600 by using EROS, which would require the use of full-head montages for both measures and computation of forward models for

the ERP data. Further, there may not be complete correspondence of the ERP and EROS effects because they may measure only partially overlapping brain activity.

The syntactic and semantic anomalies occurred in different positions in our sentences, with the semantic anomalies and their corresponding controls always in final position and the syntactic anomalies earlier. In this first EROS study of language processing, manipulations were chosen to maximize the N400 and P600 ERP effects as well as the EROS effects. Software constraints dictated an eight-word maximum sentence length, so we placed semantically anomalous words in sentence-final position to build up sufficient semantic constraint. The syntactic violations, however, did not lend themselves to sentence-final position. To estimate how much of the difference between conditions was due to word position, we analyzed the EROS elicited by midsentence semantic anomalies in 48 filler sentences. The same temporospatial pattern of S/MTC and IFC activity was found for both sentence-final and midsentence anomalies, but both the EROS and corresponding ERP effects were considerably weaker in midsentence positions. This finding suggests that the differences between semantic and syntactic anomalies are not entirely due to word position. However, the more prominent temporal-frontal oscillatory pattern, or re-entry activities, observed for the sentence-final semantic anomalies, which we interpret as the activity of an integration-prediction network in response to an anomaly, may be partly due to their position, perhaps because of sentence wrap-up processes, the need to make the decision leading to the behavioral response, or both. Similarly, we cannot make any claims about the relative speed of semantic and syntactic processing. The apparent difference in the earliest significant effect of semantic versus syntactic deviance may be due to factors such as the degree of contextual constraint or magnitude of the deviance.

We recorded activity only from the left hemisphere. However, other investigators (Just, Carpenter, Keller, Eddy, & Thulborn, 1996) have suggested that the right S/MTC and IFC also are involved in sentence comprehension. The unexplained variance in the EROS–ERP prediction pattern may be related in part to right-hemisphere contributions to language processing, as well as to other brain regions not measured in this study.

In summary, the current study used EROS to observe the temporal and spatial dynamics of activity in S/MTC and IFC in processing semantic and syntactic anomalies, confirming previous observations of the importance of these regions in sentence processing. The data suggest a flow of activation from posterior (temporal) to anterior (frontal) regions, as well as the involvement of distinct areas in the temporal cortex in the processing of semantic and syntactic anomaly and overlapping areas in the frontal cortex. Importantly, spatially and temporally distinct components of EROS were related to the N400 and P600 ERP components elicited by semantic and syntactic anomalies. Finally, this study demonstrates the feasibility of using EROS to study language processing.

# Chapter 3

## Dissociating Representations of Stimulus Contingencies with Event-related Optical Signal (EROS) and Event-related Potentials (ERPs)

### Introduction

Humans are regarded as a statistical learning machine; we extract contingencies between objects or events based on the frequency, probability, and similarity among them. A large body of animal and human research in various fields in psychology has been conducted on how features, objects, or events are classified, e.g., conditioning, categorization (e.g., Nomura, et al., 2007; Schnyer, et al., 2009), language (e.g., Opitz & Friederici, 2004), sequence learning (e.g., Bischoff-Grethe, Goedert, Willingham, & Grafton, 2004; Grafton, Hazeltine, & Ivry, 1995), and reasoning (e.g., Wendelken, Nakhabenko, Donohue, Carter, & Bunge, 2008).

Understanding regularities in the temporal order of stimuli, i.e. their sequential structure, can help one to predict the identity of the next stimulus and improve processing efficiency (Koch & Hoffmann, 2000). Cleeremans (1991) proposed a connectionist model sensitive to the regularity in the temporal sequence of stimuli. In this model, an input interacts with an existing context unit, formulated based on a previous input, to produce a response. Elman (1990) showed that this model can be applied to detect the grammar or syntax governing the letter sequence in words, or word sequence in sentences. The same model was found to underlie the procedural memory mechanism which supports both repetition priming and skill learning (Gupta & Cohen, 2002). These authors proposed that the ability to integrate temporal structures across time requires continuous tuning of an internal representation or processor for optimum performance,

although different representations or processors may be involved in repetition priming and skill learning.

In order to understand the rules governing the regularity of a sequence of stimuli, similarity in the temporal relationship between stimuli needs to be detected. Such temporal relationships need to be categorized under one or several classification schemes, which we label as contingency rules. These contingency rules may vary in the processing requirements, meaning that different mechanisms may be involved in tuning the internal representations for optimum performance. Rules may require formulation of a static internal representation, a dynamic internal representation, or both. We conceptualize a static internal representation as a template. The most common pattern of stimuli establishes this template, and once the template is learnt, it can be used for predicting the next incoming stimuli. On the other hand, a dynamic representation performs a running update of a model or template to predict future events for optimum performance. For example, if the rule specifies that 20% of events are type A, constant updates of the dynamic representation are needed to see whether the rule is met.

The way in which our brain represents regularities in stimulus sequences may depend on the complexity of these processing requirements of the rules. Rules requiring the sole application of static templates may involve a simple circuitry and can be represented at a level that is proximal to the sensory input – in psychological terms, they may be sustained at a sensory memory level. Rules requiring dynamic representations may involve a more complex circuitry, in which information is maintained over a longer stimulus sequence and irrelevant information is ignored. The processing of these rules may use forms of representations that are more distal from the sensory input – in psychological terms, they may be sustained at a working memory level. In

short, representations that differ in the complexity of processing requirements are likely to be sustained at different levels within the brain.

Here we investigate this hypothesis by contrasting the brain activity associated with the representation of contingency rules with different processing requirements. Our predictions are that as the processing requirements become more complex, the latency of brain activity associated with their representation will increase, and the location of the brain response will become more distal from the sensory input (e.g., from secondary sensory cortex, such as the superior temporal cortex – STC -- to higher level multimodal cortex, such as the inferior frontal cortex -- IFC). To investigate this hypothesis we use measures of brain activity combining temporal and spatial resolution: specifically, the event-related brain potential (ERP) and the event-related optical signal (EROS).

To date, most of the brain imaging studies using the oddball paradigm have focused on studying deviance detection; however, it is equally important to understand the extraction of regularity among a sequence of stimuli. Consistent with the premises of our work, research on regularity detection and rule learning has emphasized two different forms of ERP activities and brain regions as relevant: the Mismatch Negativity (MMN) (Näätänen & Michie, 1979; Näätänen, Paavilainen, Rinne, & Alho, 2007) (MMN), an ERP component which has been associated with secondary auditory areas, and the P300, an ERP activity which has been associated with attentional and working memory and with fronto-parietal brain circuits component (Donchin, 1981; Fabiani, Gratton, & Federmeir, 2006).

The MMN is a very extensively researched ERP component, typically elicited by rarely occurring auditory stimuli with a latency between 150 and 250 ms from stimulation. Recent

work has emphasized how the MMN can be used as a deviance detector to investigate how the regularity in a sequence is represented in the brain. For example, Horvath, et al. (2001) showed that multiple representations of contingency rules can be simultaneously maintained in the absence of attention. Specifically, a train of stimuli governed by two basic temporal rules (Koch & Hoffmann, 2000), a statistical and a relational rule, were presented. The statistical rule indicated the overall probability of a particular stimulus to occur, e.g., 80% of stimuli were A, and the remaining 20% were B. The relational rule indicated the systematic relations among consecutive stimuli, e.g., an alternation sequence with two types of stimuli (ABABAB…). Alternatively, this relational rule could be understood as a statistical rule in a much smaller temporal scale i.e. all succeeding stimuli must be different from the previous stimuli in an alternating sequence. MMN responses were found when either or both rules were violated, indicating the coexistence of representations of both rules. van Zuijen et al. (2005) found individual differences in the recruitment of abstract rules. Musicians were sensitive to multiple abstract rules governing the temporal relations among tones in a series, while non-musicians were only sensitive to one of the rules.

The P3 component (Donchin, 1981; Fabiani, Gratton, & Federmeir, 2006) has been widely applied to study stimulus probability (e.g. Brumback, Low, Gratton, & Fabiani, 2005). Duncan-Johnson and Donchin (1977) showed that the amplitude of the P300 was inversely correlated to stimulus probability. The amplitude of the P300 also reflects local probability changes in the stimulus sequence (Brumback, Low, Gratton, & Fabiani, 2005; Squires, Petuchowski, Wickens, & Donchin, 1977). The P300 was suggested to reflect context updating of working memory that affects subsequent responses (Donchin, 1981; Donchin & Coles, 1988). Lesion studies(Knight, 1984) (Knight, Scabini, Woods, & Clayworth, 1989), fMRI studies

(Kirino, Belger, Goldman-Rakic, & McCarthy, 2000; Linden, 2005; Polich, 2007), and multimodal brain imaging studies with ERP and fMRI (Benar, et al., 2007; Bledowski, et al., 2004; Calhoun, Adali, Pearlson, & Kiehl, 2006; Mantini, Corbetta, Perrucci, Romani, & Del Gratta, 2009) revealed that the fronto-parietal network modulates P300. In other words, working memory involves an interaction of the prefrontal-executive and parietal-perceptual system (Fuster, 2001; Fuster, 2006; Fuster, Bodner, & Kroger, 2000).

The current study employed an approach different from the MMN and P300 studies reviewed above to investigate regularity representation in the human brain. Instead of comparing deviant against standard stimuli in an oddball paradigm, we explored the neural substrates of contingency rules by comparing the brain responses to stimuli governed by multiple rules simultaneously. Specifically, we measured the brain responses to a sequence of stimuli governed by contingency rules with different properties using event-related optical signal (EROS) and ERP. A sequence of high and low tones governed by three underlying contingency rules was presented to participants. The participants were naïve to the rules at the start of the experiment.

The three rules were a statistical (Global Probability) rule, a relational (Single Repetition) rule, and a more complex, context-related (Local Probability) rule. The Global Probability rule dictated that 75% of the tones were alternation (i.e. …ABA**B**…), while 25% were repetition (i.e. …ABA**A**…). The Single Repetition rule dictated that the tone following a repetition must be an alternation (i.e. …ABAA**B**…). The Local Probability rule dictated that repetition could only occur in odd trials. If one kept track of the tones as pairs, the first tone of the pair had a 50% chance of being an alternation or a repetition, while the second tone must be an alternation (i.e. AB **B**A BA **A**B…).

These rules differed in two dimensions, template mismatch and requirement of contextual or working memory updating. We conceptualized that in order to apply the Global Probability and the Single Repetition rule to help the processing of the current tone, a template with properties of the most probable incoming tone needed to be formed. In the Global Probability rule, because most of the tones were alternations, an alternation template was formed. 75% of the tones fitted into this tone alternation template; however, for the other 25% of tones with repetitions, the alternation template was violated. On the other hand, the template in the Single Repetition rule described the probabilistic nature of the incoming tones. It means that the incoming tone may be an alternation or repetition, and its certainty could only be established after the presentation of the tone. Most of the tones had this probabilistic nature, except for the tone right after a repetition. Under the Single Repetition rule, the sequence never contained two repetitions (i.e. …ABA**AA**…). Thus, the tone after the repetition must be an alternation and so the deterministic nature of this tone violated the probabilistic nature in the template. In contrast, no template was needed to be formed for the Local Probability rule, because its nature was deterministic: the position of the tones within a pair could predict the next incoming tone. For the memory updating dimension of the rules, we conceptualized that both the Global Probability rule and the Local Probability rules had this feature. Both of these rules required integrating the property of the current tone into the current memory for subsequent tone processing. In the Global Probability rule, the alternating/repeating nature of the current tone was integrated into memory to help keep track of the ratio of alternations to repetitions. Deviation of this ratio from that specified by the Global Probability rule (i.e. 75/25) biased the response system towards processing an incoming tone as a potential alternation or repetition, depending on what the current ratio was biased towards. Under the Local Probability rule, the ordinal position of the

current tone within a pair (odd versus even) was integrated into memory to inform the ordinal position of the next tone. Because repetitions can only occur in odd trials while even trials must be alternations, knowing the ordinal position of a tone could reduce the work load of the response system. By contrast, for the Single Repetition rule, no memory updating was required, because the sequence started afresh after a repetition and did not depend on previous stimuli in the sequence. We predicted that different rules are represented in distinct locations in the brain and/or at distinct moments in time according to these two properties.

## Methods

### Participants

Sixteen participants (age range 19-27) who were naïve to the contingency rules at the beginning of the experiment entered the study after giving informed consent. All of the procedures were approved by the Institutional Review Board (IRB) of the University of Illinois at Urbana-Champaign. According to self-report, all participants had normal hearing, were not taking any psychoactive medications, and had no history of neurological disorders or head trauma. Scores on the Edinburgh Handedness Inventory (Oldfield, 1971) showed that all of the participants were right-handed.

### Stimuli and procedure

A forced choice reaction time task was used in this study. Participants were presented with auditory tones (70 dB sound pressure level) of 500Hz (high frequency tone) and 350Hz (low frequency tone), and were required to respond to the high/low tones by pressing the left/right buttons on the response box with their left/right thumbs. The mapping of high/low tones to response hands were counterbalanced across participants. Throughout the experiment,

participants were asked to fixate on a cross at the center of the computer screen. The duration of all tones was 400 ms and the interstimulus interval was 1600 ms. The interval allocate for subject responding consisted of the first 1200 ms after tone offset.

A sequence of high and low tones was presented to the participants. In each experimental block, 50% of the tones were high tones and the other 50% were low tones. However, among successive tones, 75% followed an alternating pattern, i.e. high-low or low-high, and 25% followed a repetition pattern, i.e. high-high or low-low. This ratio of alternation versus repetition was regarded as the Global Probability (or "statistical") rule. In addition, the sequence was set up such that when a repetition occurred, the tone that followed would always be an alternation (i.e., there were never three identical tones in a row). This was regarded as the Single Repetition (or "relational") rule. Finally, the tones could also be thought of as occurring in pairs, with odd trials having a 50/50 chance of being high or low, and even trials always occurring as an alternation and therefore 100% predictable given the preceding tone. To use the predictive information, some representation of the odd/even pairing must be available. The pairing was not explicit, because the tones had a constant stimulus onset asynchrony and participants were not informed of this aspect of the contingency, and therefore recognition of the presence of this rule require keeping track of the context in which stimuli were presented. The predictability based on tone pairs was regarded as the Local Probability (or "context-related") rule. The first two tones of each block were omitted in the following analyses.

Based on these contingency rules, the tones were categorized into four types i.e. odd trial of an alternation tone pair (T1Alt), even trial of an alternation tone pair (T2Alt), odd trial of a repetition tone pair (T1Rep), and even trial of a repetition tone pair (T2Rep), independent of the frequency of the tones i.e. high or low tones (Figure 3.1). Each odd trial (T1) and the following

even trial (T2) were regarded as a tone pair. Tone pairs were classified into alternation (Alt) and repetition tone pairs (Rep), based on the relationship between the odd trial of the current pair and the even trial of the previous tone pair. An Alt pair was one where the odd trial of the current tone pair was different from the even trial of the previous tone pair, while a Rep pair was one where the odd trial of the current tone pair was the same as the even trial of the previous tone pair. Due to the Local Probability rule, even trials must be alternating, i.e. it must be an alternation within any tone pair.

ERP and EROS data were recorded in four separate sessions (two ERP sessions and two EROS sessions) in the following sequence: 1) first ERP session, 2) first EROS session, 3) second ERP/EROS session, 4) second ERP/ EROS session. In the ERP sessions, 98 tones (50% high and 50% low tones) were presented in each experimental block. There were five experimental blocks in each ERP session (i.e. 490 tones were presented per ERP session). In the EROS sessions, 30 tones (50% high and 50% low tones) were presented in each experimental block. There were ten experimental blocks for each EROS recording montage (i.e., light source/detector configurations). In total, four recording montages were collected across two sessions, resulting in 1200 tones being presented during the EROS sessions. As learning of the contingency rules (which were not explicitly explained to the subjects) was likely to occur during the first ERP session, the data from the two ERP sessions were analyzed separately.

**EROS recording and preprocessing**

Optical data were recorded from two frequency-domain oximeters (Imagent, ISS Inc., Champaign, Illinois, USA). Frequency-modulated near infrared light (830 nm, modulated at 110MHz) emitted from laser diodes was channeled to the participant's head via individual

optical fibers (diameter = 400 μm). Light that scattered through the head was collected by optical fiber bundles (diameter = 3 mm) connected to photo-multiplier tubes (PMTs) used for the measurements. The PMTs were modulated at 110.00625MHz, generating a 6250 Hz heterodyning frequency (i.e. cross-correlation frequency). The A-D sampling rate was 50 kHz and the output current was Fast Fourier Transformed (FFT) and relative phase delay measures (in picoseconds) were computed.

The light source and detector fibers were held on the participants' head using a modified motorcycle helmet. In each EROS session, optical data was collected from two montages, each with 192 channels (i.e., 8 detectors and 12 time-multiplexed light sources from each of 2 Imagents). The final sampling frequency was 52.0833 Hz (19.2 ms to cycle through the 12 time-multiplexed light sources). The spatial position of each montage was such that, when combined, the four montages provided coverage for most of the cortical surface (see Figure 3.2). The order of the montages was counterbalanced across participants. Source-detector distances ranged between 10 and 65 mm, but channels with distances less than 15 or greater than 55 mm were excluded from further analysis (Gratton, et al., 2006).

The locations of sources and detectors, as well as the nasion and pre-auricular points of each participant were digitized with a Polhemus Fastrak 3Space 3D digitizer (Colchester, Vermont, USA). Volumetric T-1 weighted magnetic resonance images (MRI) were also obtained for each participant using a Siemens Magnetom Allegra 3 Tesla scanner. The nasion and pre-auricular points were marked with Beekley Spots (Beekley Corporation, Bristol, CT) in the MRI scan for co-registration of the functional optical data with the structural MRI using the procedure described by Whalen and colleagues (Whalen, Maclin, Fabiani, & Gratton, 2008). The

individually co-registered data were then Talairach transformed to permit registration across subjects (Talairach & Tournoux, 1988).

The optical data were corrected off-line for phase wrapping, normalized to a mean of zero, pulse corrected (Gratton & Corballis, 1995), and filtered with a 1-10 Hz band-pass filter. The data were then divided into epochs around each tone with 200 ms pre-stimulus baseline and 800 ms post-stimulus recording. Finally, these epochs were averaged for each channel, time point, and tone type. The influence of remaining noisy channels was reduced by eliminating any channels that had standard deviations of phase greater than 160 picoseconds from the analysis.

The averaged data were analyzed using in-house software, *Opt-3D* (Gratton, 2000). The optical signal for a given voxel was defined by combining channels whose mean diffusion paths (modeled as a curved ellipsoid) intersected for a given brain volume voxel (Wolf, et al., 2000). An 8-mm Gaussian filter (based on 2 cm kernel) was used to spatially filter the data. The group level *t*-statistics were calculated using an error term pooled across time. These *t*-values were then converted to *Z*-scores and orthogonally projected onto images of the superior and sagittal surfaces of a brain in Talairach space.

**ERP recording and preprocessing**

EEG was recorded with Ag/AgCl electrodes at 21 scalp locations based on the standard 10/20 system. All were referenced to an electrode placed on the left mastoid. Four electrodes, one above and one below the right eye and two at the outer canthi of each eye were used for bipolar vertical and horizontal EOG recordings. The EEG was filtered online using a 0.01 to 30 Hz band pass and sampled at 100 Hz. Electrode impedance was below 10 kOhms. The EEG data were divided into epochs around each tone with 200 ms pre-stimulus baseline and 1000 ms

post-stimulus recording. Ocular artifacts were corrected (Gratton, Coles, & Donchin, 1983) and trials with voltage changes greater than 200 μV across the 1200-ms recording window were rejected. This resulted in the rejection of ~15% of the trials. Similar to the EROS data, the electrophysiological data were signal-averaged for each subject, channel, and tone type with time-locking to the onset of the tones.

**Data analysis**

All statistical analyses (for behavior, ERPs, and EROS) presented in this paper are based on the use of planned contrasts referring to sensitivity to each of the three rules. The experimental design allowed for three orthogonal contrasts, one for each rule, to be set up. One sample t-tests (one-tailed) were used to test for the significance of individual contrasts, demonstrating the sensitivity of each dependent variable to each rule (directional contrasts were used because, for each dependent variable, specific hypotheses about the direction of the effect could be made). It should be noted that as the weights of the contrasts were not normalized, statistical tests comparing the difference between contrasts are invalid. Different weights were assigned to the four trial types according to the contingency rules (Figure 3.1) with the constraint that the sum of the cross products of the contrasts was equal to zero (i.e., orthogonal contrasts). The weights assigned can be interpreted as the relative effort needed for processing the tones. It was expected that tones with positive weights were associated with longer response times and stronger brain responses compared to tones with negative weights in a particular contrast. Zero weights were assigned to tones that were not involved in the calculation of a particular contrast.

Under the Global Probability (statistical) rule, the weights assigned to T1Rep, T1Alt, T2Alt and T2Rep were 3, -1, -1, and -1, respectively. The processing of the repeated tone was different from that of the alternated tone, as tone repetition violates the alternation template.

Furthermore, a repeating tone also required memory update in order to keep track of the ratio of repetition to alternation to assist future tone processing. It was predicted that the rare repetition tone would lead to longer response times and stronger brain responses.

Under the Single Repetition (relational) rule, the weights assigned to T1Rep, T1Alt, T2Alt and T2Rep were 0, 1, 1, and -2 respectively. In this contrast, only alternation tones were compared, which means T1Rep was excluded. As mentioned above, the tone (i.e. T2Rep) after a repetition (i.e. T1Rep) must be an alternation. That means that the deterministic nature of T2Rep was different from the probabilistic nature of other tones (i.e. T1Alt and T2Alt) as one can only be certain of their alternation/repetition nature after the tones were presented. Thus, we predicted that the processing of T2Rep would be easier and faster than that of alternation tones, i.e. T1Alt and T2Alt.

Under the Local Probability (context-related) rule, the weights for T1Rep, T1Alt, T2Alt and T2Rep were 0, 1, -1, and 0 respectively. This contrast compared the first tone (T1Alt) against the second tone (T2Alt) of the alternating pair. The repetition tone pair (T1Rep and T2Rep) was excluded in contrast because this pair was covered by the previous two contrasts. Under this rule, if one could keep track of the identity of the tone pair, processing of tone 2 would be easier as it must be an alternation. In other words, comparing T1Alt and T2Alt, the reaction to T1Alt should be slower and the processing of T1Alt should be more effortful as it required integrating information about the ordinal position of the current tone.

## Results

### Behavioral results

Mean reaction times across participants and standard error of the mean of the four trial types for the four ERP/EROS sessions are shown in Figure 3.3. The reaction time in the first ERP session, i.e. the first recording session of the entire experiment, was much longer than in the other three sessions for all four trial types, indicating that learning was still occurring during this session. Figure 3.3 also showed differences between trial types, indicating learning of the contingency rules. To better understand which specific rule was learned, we used orthogonal contrasts for each rule (see Methods section). Figure 3.4 showed the mean value and standard error of mean for the contrast analysis on reaction time for each of the three rules. Repeated measures ANOVA showed a significant contrast by ERP session interaction effect, $F(2,14) = 7.35$, $p < .05$: Differences in the responses to different rules between the first and the second ERP sessions were found, confirming that learning had occurred in between them. Therefore the first ERP session was excluded in the following analyses. A similar analysis was conducted for the EROS session, but no reliable contrast by EROS session interaction effect was found, $F(2,14) = 2.28$, $p > .05$, suggesting that no significant rule learning occurred in between these two sessions. Therefore the two EROS sessions were collapsed to increase the statistical power of the analysis. Thus, the following analyses focused on the second ERP session and the combined EROS session. Table 3.1 summarized the statistical results for the contrast analysis on reaction time. Contrasts corresponding to all of the three rules were statistically significant, for the second ERP and both EROS sessions, but not in the first ERP session. These data indicate that, after the learning phase occurring in the first ERP session, the three presentation rules were utilized to process the tones in the second ERP session and both EROS sessions.

**ERP results**

The ERP waveforms of each trial type at Fz, Cz, and Pz electrodes are shown in Figure 3.5. The waveforms showed differential responses at 200-250ms (corresponding to the N2 component – which is analogous to the MMN in active tasks) at Cz and 250-450ms (corresponding to the P300 or P3 component) at Pz to different trial types. Contrast weights were applied to reconstruct waveforms corresponding to the application of contingency rules (Figure 3.6) and time windows were set up to calculate the mean amplitude of the N2 effect (between 200 and 250 ms post-stimulus) and P3 effect (between 250 and 450 ms post-stimulus) for each contrast. Contrast analysis showed significant responses in the N2 time window for the Global Probability (statistical) and Single Repetition (relational) rules, but not the Local Probability (context-related) rule, and significant responses in the P3 time window for the Global Probability (statistical) and Local Probability (context-related) rules, but not the Single Repetition (relational) rule, (Table 3.2). In other words, ERP responses corresponding to all three rules were found. However, they were processed at different times along the information processing stream. Processing related to the Global Probability (statistical) rule started at around 200 ms and sustained through 400ms, while processing of the Single Repetition (relational) rule showed an earlier onset than that of the Local Probability (context-related) rule.

**EROS results**

Statistical maps of EROS data based on the same contrast approach presented above are shown in Figure 3.7. Note that different maps can be obtained for different latencies. Consistent with the ERP findings, the EROS results showed bilateral temporal activities sensitive to the Global Probability (statistical) and Single Repetition (relational) rules but not to the Local Probability (context-related) rule at latencies of 211ms-287ms. Again consistent with the ERP findings, frontal activities sensitive to the Global (statistical) and Local Probability (context-

related) rules but not to the Single Repetition (relational) rule were found at 307ms-460ms latencies. A left temporal area sensitive to the processing of the Global Probability (statistical) and Single Repetition (relational) rules at an earlier time interval was also sensitive to the Local Probability (context-related) rule at 345ms. Consistent with our predictions, these results showed that the simpler rules, the Global Probability (statistical) and Single Repetition (relational) rules, were processed in temporal regions at earlier time intervals while the earliest evidence for recognition of the more complex rule, Local Probability (context-related) rule, occurred within frontal parietal network at later time intervals.

**Discussion**

The current study employed a novel procedure to explore the brain responses to stimuli governed by multiple rules simultaneously. Instead of comparing the deviant against the standard stimuli, orthogonal contrasts were used to investigate representation in the human brain of sequential regularities embedded within a 50/50 choice reaction time task. Behavioral results suggested that learning occurred across the recording sessions. In the first ERP session, the regularity in the tone sequence was not well established. However, all of the contingency rules, Global Probability (statistical), Single Repetition (relational), and Local Probability (context-related) rules, were utilized in the last three recording sessions for processing the tones. Although we did not systematically ask the participants about the presence of each contingency rule, they typically expressed awareness of the presence of these regularities by the end of first ERP session. It should be noted that there were differences between the responses elicited in the ERP and EROS sessions. The reaction time in both ERP sessions (even the second one) was slower in general. This probably reflects the fact that the longer experimental block in the ERP sessions made the task more difficult. This may particularly affect application of the Local Probability

(context-related) rule, as it may lead to increased difficulty in keeping track of the positions of individual trials within the block.

The ERP results showed that the rules were processed at different latencies. Evidence for stimulus differentiation based on the Global Probability (statistical) rule emerged at200 to 400 ms post-stimulus, for the Single Repetition (relational) rule at around 200ms, and for the Local Probability (context-related) rule at around 400ms. In addition to confirming the latency differences shown by ERP, the EROS results further suggested that the brain locations exhibiting this differentiation varied for different rules. For the Global Probability (statistical) rule, differentiation was evident first in the superior temporal cortex, followed by the prefrontal cortex. For the Single Repetition (relational) rule, differentiation was first evident in the temporal region, and for the Local Probability (context-related) rule, the first evidence of differentiation occurred in the frontal region. This suggested temporal and spatial differentiation in the processing of the contingency rules.

The ERP and EROS results are consistent with the use of the template mismatch and contextual or working memory updating dimensions to classify contingency rules suggested in the Introduction. The concepts of template mismatch and context-related memory updating also fit well under the framework of ERP N2 and P3 effects, respectively. It has been suggested that the N2 is related to mismatch processing when incoming stimuli are not consistent with the expectancy derived from the memory template (Folstein & VanPetten, 2008), and the amplitude of N2 increases with decreases in the probability of the stimuli (Bruin & Wijers, 2002). The P3 component (Donchin, 1981; Fabiani, Gratton, & Federmeir, 2006) is suggested to reflect context updating of working memory that affects subsequent responses (Donchin, 1981; Donchin &

Coles, 1988), and its amplitude reflects local probability changes in the stimulus sequence (Brumback, Low, Gratton, & Fabiani, 2005; Squires, Petuchowski, Wickens, & Donchin, 1977).

In the Global probability (statistical) rule, it is not surprising that a standard template was formed for the alternating pattern, as 75% of the trials were alternations. The violation of the alternating pattern, i.e. repetition, produced template mismatch as indicated by the N2 effect and reflected as early temporal activities with EROS. In addition, the Global Probability (statistical) rule also required constant memory updating of the ratio of alternation to repetition, so that appropriate responses could be made to future stimuli. This updating process may then result in the elicitation of an enhanced P300. This result is consistent with previous working memory studies of the fronto-parietal network, suggesting that the frontal cortex integrates information across time (Fuster, 2001; Fuster, 2004; Fuster, 2006). The spatiotemporal properties of the observed frontal activity were similar to those found in a previous EROS study that used auditory active and passive oddball paradigms (Low, Leaver, Kramer, Fabiani, & Gratton, 2006).

For the Single Repetition (relational) rule, the template for incoming tones was probabilistic in nature; however, the tone after the repetition (T2rep) was deterministic in nature and so caused a mismatch to the template, which elicited the ERP N2 and EROS temporal response. Because the tone that followed a T2rep was not contingent upon T2rep, no updating of the memory template was needed, and so no P3 or associated frontal cortex activity was observed.

In the Local Probability (context-related) rule, keeping track of the ordinal position of the tone pair could reduce the work load of the response system, because repetitions could only occur in odd trials while even trials must be alternations. The memory updating process in the

Local Probability (context-related) rule was associated with the observed P3 effect and EROS frontal activity. Note, however, that an N2 and activity in superior temporal cortex were not observed in association with this rule. This may reflect the greater representational demands underlying this contingency rule: This rule requires integrating ongoing information with contextual information (related to the ordinal position of the stimulus within the sequence), which may simply not be representable at early (proximal to sensory input) cortical levels, but can only be achieved at more distal representational level (see Fuster, 2001; Fuster, 2004; Fuster, 2006 for an example of a hierarchical view of processing within the brain).

The current study provided evidence that the brain is sensitive to multiple contingency rules simultaneously. The contingency rules can be differentiated from each other spatiotemporally along the template mismatch and memory update dimensions. More interestingly, the results showed that despite the complexity of the three contingency rules, our brain can decompose the mechanisms required in each contingency rule into smaller units and apply basic cognitive tools (i.e. template comparison and memory updating) to handle the rules.

This conclusion can also be understood under the framework of the connectionist model proposed by Cleeremans and McClelland (1991) in which an input interacts with an existing context unit formulated based on a previous input to produce a response. The template and the memory updated can be conceived as the context unit in the connectionist model. In addition, the current study suggested that there are different types of context units involved, depending on the processing requirements of the rules. This idea is also analogous to Gupta and Cohen (2002)'s idea that different representations or processors are being tuned for optimum performance in repetition priming and skill learning.

There are striking similarities in the temporal components of the EROS and ERP results. The EROS temporal and frontal cortex activities were observed in similar time windows as the ERP N2 and P3 components. More interestingly, comparison between ERP P3 components elicited by the Local Probability (context-related) and Global Probability (statistical) rules showed a small difference. The P3 effect for the Global Probability (statistical) rule occurred earlier than that for the Local Probability (context-related) rule. This small temporal difference was also reflected in the slightly earlier frontal cortex activity observed for the Global Probability (statistical) rule than for the Local Probability (context-related) rule. Similar to previous studies with simultaneous EROS and ERP recordings (Tse, et al., 2007; Tse & Penney, 2008) that showed specific correlations between EROS and ERP amplitudes, the current study also suggests temporal correspondence of ERP components and EROS.

Different from the traditional approach of comparing brain responses to standards and deviants, the current study applied a statistical method, contrast analysis, to investigate the representation of contingency rules by comparing brain responses to multiple stimuli. Although the procedures in assigning contrast weights and calculating contrasts were more complicated than the simple one-to-one contrast used to compare standards to deviants, our design reduced the complexity and length of the recording procedure, as multiple manipulations and different recording blocks would have been needed had we chosen the traditional approach. However, the contrast analysis approach required specific predictions on how the contingency rules differed from each other in order to test them systematically.

In summary, the brain can simultaneously hold different models of stimulus sequences within the information processing stream; however, these representations can be differentiated

from each other, both in terms of latencies and locations of activations, using both ERPs and EROS, based on their complexity and type of processes required for their instantiation.

# Chapter 4

## Read my Lips? Dissociating Audiovisual Integration and Deviance Detection with Event-related Optical Signals (EROS)

### Introduction

Perception of external stimuli not only relies on the sensation from individual modalities but also depends on the integration across multiple senses. Audiovisual speech perception in noisy environments is one of the most common examples in our everyday life that requires integration across modalities. As auditory speech is produced by changing the vocal tract configuration (mouth opening, closure, shape, place of articulation, and manner of articulation), it is not surprising that seeing the lip or face movements of the speaker can improve the intelligibility of speech, especially in noisy environments (Ross, Saint-Amour, Leavitt, Javitt, & Foxe, 2007; Sams, Mottonen, & Sihvonen, 2005; Sumby & Pollack, 1954; Winkler, Horvath, Weisz, & Trejo, 2009).

More interestingly, seeing the lip movements of the speaker can actually modify the perception of a speech sound and produce an auditory illusion, the McGurk effect (McGurk & Macdonald, 1976). For example, the syllable /ba/ presented auditorily is perceived as /da/ when coupled with a face presenting the lip movements associated with the syllable /ga/. This indicates that the perception of bimodal stimuli involves audiovisual integration. The McGurk effect provides a basis for investigating audiovisual integration and distinguishing the unimodal and multimodal sensory cortices.

The McGurk effect is an automatic process (Soto-Faraco, Navarra, & Alsius, 2004): Participants perceive the illusion independent of their awareness of the incongruent stimuli. The McGurk effect can be obtained even when the audio and visual stimuli are not presented in

synchrony (van Wassenhove, Grant, & Poeppel, 2007). By varying the lag between the audio track, /ba/, and the visual track, /ga/, of a video, it was revealed that audiovisual integration can tolerate asynchrony with a window of 200 ms, which is similar to the temporal window of integration for acoustic stimuli (Yabe, Tervaniemi, Reinikainen, & Näätänen, 1997; Yabe, et al., 1998). More importantly, the fused percept of /da/ is biased towards a video lead (van Wassenhove, Grant, & Poeppel, 2007), suggesting that visual speech is used for predicting auditory speech. The McGurk illusion is probably elicited based on our language experience (Bulkin & Groh, 2006).

Sams et al. (1991) conducted the very first brain imaging study on the McGurk effect by using magnetoencephalography (MEG). Frequent audio-visual congruent (audio /pa/ visual /pa/) and rare incongruent videos (audio /pa/ visual /ka/) were presented while participants counted the total number of auditory stimuli. The results showed a mismatch response between the standard and deviant stimuli from 200ms to 500ms after speech onset over the left temporal area. This result showed that even when the same auditory /pa/ was used for the standard and deviant stimuli, the auditory /pa/ of the deviants is perceived differently from that of the standards. The same mismatch response was observed when the frequencies of the congruent and incongruent videos were reversed in separate blocks. These results suggested that visual information affects speech processing in the auditory cortex. The electric counter part of the magnetic mismatch response in Sams et al. (1991) was demonstrated by Colin et al. (2002) with audiovisual speech of voiced consonants (AV/bi/ as standard, A/bi/V/gi/ as deviant) and Colin et al. (2004) with voiceless consonants (i.e., /pi/ /ki/).

The passive oddball paradigm was applied to study audiovisual speech perception and the mismatch responses were believed to be the audiovisual analogue of the mismatch negativity

(MMN; Näätänen & Michie, 1979; Näätänen, Paavilainen, Rinne, & Alho, 2007). The MMN has been widely found in the auditory domain (aMMN) and is typically regarded as a negative wave occurring between 150ms to 200ms after the onset of the deviant, most visible at the Fz electrode. Recent accounts of the MMN suggested that it reflects violations of regularity or expectancy (Schröger, 2007; Winkler, 2007). MMN-like responses have also been found in other sensory modalities, e.g. the visual MMN (Czigler, 2007; Kimura, Katayama, & Murohashi, 2005; Kimura, Katayama, Ohira, & Schröger, 2009) and the somatosensory MMN (Kekoni, et al., 1997). Although the aMMN is regarded as a tool for probing sensory memory (Schröger, 2007), the memory representations involved in MMN are not necessarily unimodal but can be multimodal (Winkler, Horvath, Weisz, & Trejo, 2009).

Although MMN-like responses were commonly found (e.g., Colin, Radeau, Soquet, & Deltenre, 2004; Colin, et al., 2002; Sams, et al., 1991), these responses may reflect visual differences between the deviant and the standard, in addition to violation of regularity. Other studies (Mottonen, Krause, Tiippana, & Sams, 2002; Saint-Amour, De Sanctis, Molholm, Ritter, & Foxe, 2007) included a visual control condition to subtract out the visual mismatch and differential mouth movements between the standard and deviant, i.e. (AV deviant - AV standard) - (V deviant - V standard). This contrast removed both the pure auditory and visual mismatch responses, but the resulting waveform measured the AV interaction in addition to the pure AV deviance detection process (avMMN). The AV interaction reflects AV integration processes, including conflict resolution between auditory and visual information and increased difficulty in AV integration, which may in fact be the center of interest in studying AV speech perception with the passive oddball paradigm. However, the mixture of the AV integration and avMMN effects limited the interpretation of these results.

Despite of the drawback mentioned above, using the passive oddball paradigm to study AV integration in speech perception also has its advantage. Both active and passive paradigms have been applied to study AV integration (e.g., Callan, et al., 2003; Mottonen, Krause, Tiippana, & Sams, 2002; Saint-Amour, De Sanctis, Molholm, Ritter, & Foxe, 2007; Skipper, Nusbaum, & Small, 2005; Skipper, van Wassenhove, Nusbaum, & Small, 2007). Both paradigms are based an additive-factor logic (Besle, Bertrand, & Giard, 2009). The additive factor logic states that the brain response to a bimodal stimulus is equal to the sum of the responses to unimodal stimuli, plus an eventual interaction term reflecting multimodal processing. In other words, the interaction effect corresponds to the AV integration process. Based on the postulate of "pure insertion" inherent to the additive factor logic, the additive model also assumes that attention loads, task demands, and anticipatory effects (and all other processes not explicitly manipulated) are the same in the bimodal and unimodal conditions. However, it is not easy to equate attention loads and task demands with an active response paradigm (Besle, Bertrand, & Giard, 2009). In addition, active responses to McGurk stimuli may involve conflict resolution during stimulus-response mapping, which activates similar regions in the frontal cortex (Koechlin, Ody, & Kouneiher, 2003) that are involved in audiovisual speech perception. In order to prevent contamination of brain responses from variations in attention and stimulus-response mapping while studying the perceptual system, several investigators have proposed using a passive paradigm (e.g., Saint-Amour, De Sanctis, Molholm, Ritter, & Foxe, 2007; Sams, et al., 1991; Skipper, Nusbaum, & Small, 2005; Skipper, van Wassenhove, Nusbaum, & Small, 2007).

Separation of AV integration, AV deviance detection and possibly their interaction effect is not easy, as these processes may involve a similar frontal temporal network. ERP and MEG studies with source localization analysis suggested temporal (Mottonen, Krause, Tiippana, &

Sams, 2002; Saint-Amour, De Sanctis, Molholm, Ritter, & Foxe, 2007) and frontal generators (Saint-Amour, De Sanctis, Molholm, Ritter, & Foxe, 2007) for the mismatch response elicited in the AV passive oddball paradigm. Moreover, frontal and temporal generators have been typically found in auditory MMN studies using similar methods as well as fast optical imaging (Opitz, Rinne, Mecklinger, Cramon, & Schröger, 2002; Rinne, Degerman, & Alho, 2005; Tse & Penney, 2007; Tse & Penney, 2008; see Deouell, 2007 for review; Tse, Tien, & Penney, 2006). AV integration studies with functional magnetic resonance imaging (fMRI) demonstrated that auditory cortex is activated by visual stimuli. Calvert et al. (1997) showed superior temporal gyrus (STG) activity during silent lip reading, and the same brain region was activated in audiovisual speech perception. Bulkin et al. (2006) suggested that the STG was activated by silent visual speech when participants tried to interpret its meaning. In addition to the auditory cortex, the STG, and the superior temporal sulcus (STS), frontal cortex, i.e. the Broca's area, the premotor cortex (PMC), the supplementary motor area (SMA), and the insula, have also been found to be involved in audiovisual speech perception (Callan, et al., 2003; Miller & D'Esposito, 2005; Skipper, Nusbaum, & Small, 2005; Skipper, van Wassenhove, Nusbaum, & Small, 2007).

In the current study, we introduced a set of control conditions that allowed the dissociation of brain responses related to stimulus deviance detection from those of AV integration. Specifically, a passive oddball paradigm consisting of three block types, AV, visual (V), and AV control, was used. The AV block consisted of AV speech /ba/ as the standard, and McGurk stimuli, A/ba/ V/ga/, as the deviant. The V block was similar to the AV block except it included only the video track. As discussed above, the comparison between the AV incongruent deviant and AV congruent standard, with their visual differences eliminated, reflected a mixed response of AV integration and AV deviance detection (mixed response contrast). The AV

control block, based on the control condition suggested by Jacobsen and Schröger (2001), was included to dissociate the avMMN from AV integration. Six types of randomly presented AV speech, including the McGurk stimuli, were included in the AV control block. The probability of each of the six types of speech was matched with that of the AV deviant, so that the McGurk stimuli in the AV control block was identical to that in the AV block except for the absence of a contextual effect built up by preceding stimuli. The repeating standard in the AV block established regularity, while such regularity was absent in the AV control block. Comparison of the AV deviant in the AV block to the McGurk stimuli in the AV control block (avMMN contrast) allowed us to isolate a "pure" avMMN effect. Note that previous studies suggest that the McGurk illusion is largely independent of context: Although perception of the McGurk stimuli may be affected by preceding stimuli, such influence is limited to the lexical context (Windmann, 2004). Some studies even suggested that the McGurk illusion is completely automatic and is not influenced by the lexical context of preceding stimuli (e.g., Sams, Manninen, Surakka, Helin, & Katto, 1998).

Comparison of the mixed response contrast and the avMMN contrast allows dissociation of brain responses related to AV integration, AV deviance detection, and their possible interaction effect. Specifically, those brain responses that overlap between these two condition in both location and latency should be interpreted as reflecting AV deviance detection process or avMMN, while responses unique to the mixed response contrast should be considered related to the AV integration process. More interestingly, the observation that the same piece of cortex is activated in both contrasts, but with different amplitude or latency between them, can be taken to indicate the occurrence of interaction of the AV integration and AV deviance detection processes.

Catch trials of audiovisual speech with visual masks or tones were also included in all three block types. Participants were instructed to respond to the duration of the masks or tones by a button press. Different from the tasks employed in other studies, e.g., counting the total number of auditory stimuli (Sams, et al., 1991), this task required attention to both the auditory and visual channels, while participants were not required to actively process the speech information, so that the avMMN could be measured without contamination from response conflicts.

In addition to electrophysiology, we also measured the event-related optical signal (EROS) (Gratton, Corballis, Cho, Fabiani, & Hood, 1995). EROS uses near infrared light to measure the optical changes associated with neuronal responses. It is based on the measurement of the change in photon "time of flight" from the source to the detectors that occur after stimulus presentation, as a result of brain activity. The temporal resolution of EROS is in the millisecond range and its spatial resolution is in the centimeter range. It has been widely used in the measurement of brain responses in the visual (Gratton, et al., 2006; Gratton, Corballis, Cho, Fabiani, & Hood, 1995), auditory (Fabiani, Low, Wee, Sable, & Gratton, 2006; Sable, et al., 2007), and somatosensory modalities (Maclin, Low, Sable, Fabiani, & Gratton, 2004). It has also been applied to the study of brain activities associated with sensory and working memory (Fabiani, Low, Wee, Sable, & Gratton, 2006; Gratton, Fabiani, Goodman-Wood, & Desoto, 1998; Low, Leaver, Kramer, Fabiani, & Gratton, 2006; Medvedev, Kainerstorfer, Borisov, Barbour, & VanMeter, 2008), pre-attentive change detection (Rinne, et al., 1999; Tse & Penney, 2007; Tse & Penney, 2008; Tse, Tien, & Penney, 2006), executive function (Gratton, Rykhlevskaia, Wee, Leaver, & Fabiani, 2009) and language processing (Tse, et al., 2007). Measuring EROS in the current study permits the spatial and temporal separation of frontal and temporal cortical activities elicited by the AV integration and deviance detection processes.

## Methods

### Stimuli and procedure

Sixteen undergraduates and graduates (9 female, mean age = 23.7) from the University of Illinois Urbana-Champaign were recruited and gave informed consent to participate in the study. The experimental procedure was approved by the Institutional Review Board of the University of Illinois Urbana-Champaign. All participants were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971), and had no history of neurological disorders or head trauma. During the experiment, videos of a male native English speaker pronouncing a syllable were presented to the participants. Videos with the speaker pronouncing the syllables, /ba/, /ga/, /pa/, /ta/, /ǰa/ (ja), /ča/ (cha), /ka/, were recorded. The videos were sampled at a rate of 29.970 frames (440 pixels x680 pixels) per second with a total of 69 frames, i.e. 2.2977 seconds. The videos were aligned so that in all of them the onset of the sound occurred at frame 25, i.e. 832.5ms after the onset of the video. The audio and video tracks of three samples of each syllable were recombined to produce nine variations of each syllable. The audio tracks of /ba/ and video tracks of /ga/ were recombined to produce nine variations of the McGurk /da/ syllable.

There were two types of catch trial conditions (Figure 4.1). In the visual catch trial condition, a gray-color pixilated-noise mask briefly covering the mouth area of the speaker was inserted in one of the regular video tracks. In the auditory catch trial condition, , a brief tone of the same frequency as that of fundamental of the male speaker voice was inserted on one of the regular audio tracks. The first and last 10% of the tone duration were the rise and fall periods of the tone, and the maximum intensity of the tone was the same as the average intensity across syllables. The duration of the mask or tone was either 10 frames, i.e. 333ms or 30 frames, i.e. 999ms, long. The task of the participants was to judge the length of the mask/tone and made a

button press response. In order to make sure the participants attended to both the audio and visual channels of the stimuli throughout the entire video, the mask or tone could appear at the 6-th, 21-th, or 36-th frame of the video, i.e. 166.5ms 666ms 1165.5ms from onset of the video. All of the catch trials and the trial after the catch trials were not included in the analysis.

The videos used in this study consisted of three block types (Figure 4.2): the experimental blocks, the audio-visual control blocks, and the visual-only control blocks. In the experimental blocks, there were three types of videos: the standard stimuli which consisted of audio track /ba/ with visual track /ba/ (AV standard); the deviant stimuli which consisted of audio track /ba/ and visual track /ga/ (McGurk /da/; AV deviant); and the catch stimuli which were the standard or deviant stimuli with the gray color mask or tone inserted in the video. Nine out of 57 trials in each block (15.79%) were deviant stimuli; three out of 57 trials were catch stimuli (5.26%); and the remaining trials were standard stimuli. The ratio of standard versus deviant catch trials was maintained in each block, and 50% of the catch trials were visual catch trials while the remaining were audio catch trials. A pixilated still image of the male speaker was presented between trials for about 200ms, so that the stimuli onset asynchrony of the videos was 2500ms.

The audio-visual control blocks consisted of six types of videos, /pa/, /ta/, /ǰa/ (ja), /ča/ (cha), /ka/, and the McGurk /da/ (AV control). Nine trials (i.e. 15.79%) for each type of video were randomly presented with the addition of three catch trials (a total of 57 trials). The catch trials were one of the six types of videos described above, with a mask or a tone inserted in the video or audio track. Visual-only control blocks were similar to the experimental blocks, but the audio track of standard (V standard), deviant (V deviant) and catch stimuli were removed.

However, the tones were still present in the catch trials to make the participants attend to both audio and visual tracks.

A total of 42 blocks, 14 blocks for each block type, were presented to each participant. Half of the participants were presented with the following block order: experimental blocks, visual-only control blocks, audio-visual control blocks. For counterbalancing the block order effect, the other half of the participants were presented with the reverse block order: audio-visual control blocks, visual-only control blocks, experimental blocks

**ERP recording**

EEG was recorded with gold electrodes at 7 scalp locations based on the 10/20 system (Fz, Cz, Pz, C3, C4, left, and right mastoid) with reference electrodes placed at the nose tip. Four electrodes, one above and one below the right eye and two at the outer canthi of each eye were used for bipolar vertical and horizontal EOG recording. EEG was filtered online using a 0.01 to 30 Hz band pass and sampled at 200 Hz. Electrode impedance was below 5 kOhms. EEG and EOG were filtered offline with a 0.1-20 Hz band pass. EEG waveforms were segmented using 2500 ms epochs time-locked to the onsets of the speech burst, starting 200 ms before the onset of the auditory speech (baseline), and averaged according to stimulus condition. Ocular artifacts were corrected (Gratton, Coles, & Donchin, 1983) and epochs containing other EEG artifacts (i.e., pen outs or a range exceeding 200 microvolts) were removed from the analysis.

**EROS recording**

A frequency-domain oximeter (Imagent, ISS Inc., Champaign, Illinois, USA) was used produce frequency modulated (110 MHz) near infrared light (830 nm) from laser diodes; the light was channeled to the participant's scalp via fourty 2.5-meter-long plastic-clad silica optical

fibers (400 μm-diameter core). Twenty-four fiber optic detector bundles (3 mm diameter) were placed on the participant's scalp and light from the source fibers that passed through the participant's scalp, skull, and brain to reach these detectors was carried to photo multiplier tubes (PMTs) within the Imagent. The PMTs were modulated at 110.003125 MHz, thereby generating a 3.125 kHz heterodyning frequency (i.e., cross-correlation frequency). The A-D sampling rate was 50 kHz, the output current was Fast Fourier Transformed (FFT) and DC (average) intensity, AC (amplitude) intensity, and relative phase delay measures were computed every 1.6 ms; only phase delay data are reported here.

A custom-built head mount system held the light source and detector fibers in position on the participant's head. A single montage was used to interrogate both of the left STC, left IFC, and the occipital cortex at the same time (Figure 4.3). The montage comprised 24 detectors and 40 sources; however, because of the effects of large source-detector distances, each detector could receive light from only 16 sources, that were time-multiplexed, yielding a total of 384 source-detector pairs (24 detectors x 16 light sources, 830nm) and an effective sampling rate of 25.6 ms (approximately 39.1 Hz). This montage configuration allows high spatial resolution recording at the expense of area coverage, which did not cover the entire head, but only right and left occipital, left inferior parietal, left temporal and left inferior frontal cortex.

Similar to co-registration of fMRI data with structural MRI in fMRI analysis, the functional optical data were co-registered with individual structural MRI. T1 weighted 3D anatomical MRI images were obtained for each participant using a Siemens Trio 3 Tesla scanner. The nasion and pre-auricular points were marked with vitamin E pills in each MRI scan. The same fiducial points, as well as 150 points scattered around the scalp and eye socket regions were digitized in 3D space (Polhemus Fastrak 3Space, Colchester, Vermont, USA), and used for

co-registration with the MR anatomical data (see Whalen, Maclin, Fabiani, & Gratton, 2008). The locations of the recording points were transformed into Talairach space (Talairach & Tournoux, 1988) and these data were used for image reconstruction, conducted using in-house software (Opt-3D, (Gratton, 2000)).

The optical data were corrected for phase wrapping, transformed into picoseconds, normalized by subtracting the phase mean, pulse corrected (Gratton & Corballis, 1995), filtered with a 1-10 Hz band-pass filter, and then segmented (using epochs 2000 ms long, including a 204.8 ms -- 8 points – prestimulus baseline interval) and averaged for each channel, time point, and condition. The influence of noisy channels was reduced by eliminating from the analysis any channels that had standard deviations of the phase greater than 160 picoseconds and/or a source – detector distance less than 15 mm or greater than 75 mm (Gratton, et al., 2006).

The averaged data were analyzed using Opt-3D (Gratton, 2000). The optical signal for a given voxel was defined as the mean value of the channels that overlapped at that particular voxel (Wolf, et al., 2000). t-statistics of the phase data were calculated at group level for each voxel and converted to Z-scores. Statistical maps of the optical signal for each data point, i.e. a time window of 25.6 ms, were generated by projecting the Z-score onto the right lateral view with the right lateral surface of a template brain as the background and an 8 mm spatial filter was applied on the EROS data (functional data). Specifically, the EROS data were back-projected onto the lateral and posterior view, but not on the lateral and posterior surface of the brain, and averaged across the x-axis (left-right). Because EROS can only interrogate up to 2-3 cm deep from the surface of the cortex, this projection procedure is an effective way for data reduction. As we are interested on the right hemisphere and visual cortex of the brain, only data from the middle sagittal plane to the surface of the right hemisphere and data from the posterior coronal

plane were included in the analysis. Because the statistical map was surface-projected onto a template brain, the Talairach coordinates reported here comprise only y (anterior-posterior) and z (dorsal-ventral) values for the sagittal plane and x (left-right) and z values for the coronal plane.,

### EROS and ERP Analyses

In order to compare the spatiotemporal dynamics involved in AV deviance detection and AV integration, the analysis focused on two contrasts. In the first contrast (mixed response contrast), the brain response to the AV deviant was subtracted from that of the AV standard, with the difference between the V deviant and V standard removed, i.e. (AV deviant - AV standard) - (V deviant - V standard). This AV deviant versus AV standard interaction contrast showed the mixed responses of AV integration and avMMN. In the second contrast (avMMN contrast), the brain response to the AV deviant was subtracted from that of the AV control to show the avMMN response. Three ROIs, frontal, temporal, and occipital, for statistical analysis were constructed based on previous studies on deviance detection and AV integration (Callan, et al., 2003; Miller & D'Esposito, 2005; Mottonen, Krause, Tiippana, & Sams, 2002; Opitz, Rinne, Mecklinger, Cramon, & Schröger, 2002; Rinne, et al., 1999; Saint-Amour, De Sanctis, Molholm, Ritter, & Foxe, 2007; Skipper, Nusbaum, & Small, 2005; Skipper, van Wassenhove, Nusbaum, & Small, 2007; Tse & Penney, 2007; Tse & Penney, 2008; Tse, Tien, & Penney, 2006)They were analyzed independently.

## Result

### Behavioral results

Behavioral results showed that all participants responded correctly to at least 89% of the catch trials, which suggested that participants attended to both visual and auditory channels throughout the experiment.

**ERP results**

Figure 4.4 showed the averaged ERP waveforms of the mixed response contrast and the avMMN contrast. Running t-tests (against zero) revealed a significant increase in negativity from 270 ms to 375 ms in the waveform of the mixed response contrast and an earlier negativity from 230 ms to 265 ms in the waveform of the avMMN contrast. Based on the running t-test result, mean amplitude of the ERP response across an early, i.e. 230 ms to 265 ms, and a late time window, i.e. 230 ms to 265 ms, were calculated for each of the contrasts. The mean amplitude and the standard error of the mean are shown in Figure 4.5. One sample t-test against zero showed significant negativities in both the early (mean = -1.242 μV, standard deviation = 2.267, $t$ (15) = -2.191, $p < .05$) and late time windows (mean = -1.849 μV, standard deviation = 1.582, $t$(15) = -4.675, $p < .05$) in the mixed response contrast. In the avMMN contrast, significant negativity was found in the early time window (mean = -0.793 μV, standard deviation = 1.047, $t$ (15) = -3.027, $p < .05$) but not in the late time window (mean = -0.545 μV, standard deviation = 1.285, $t$ (15) = -1.696, $p > .05$).

A repeated measures ANOVA with contrast and time window as factors showed non-significant main effects of both contrast ($F$(1,15) = 2.403, $p > .05$) and time window ($F$(1,15) = 0.447, $p > .05$) but a significant interaction effect ($F$(1,15) = 4.630, $p < .05$). Analysis of simple effects with paired-sample t-tests suggested a significant difference in the mean amplitude between the contrasts ($t$(15) = -2.384, $p < .05$) in the late time window, indicating a larger negativity in the mixed response contrast. However, no significant difference between the contrasts was found in the early time window ($t$(15) = -0.694, $p > .05$). These results suggested a more sustained and larger negativity in the mixed response contrast compared to the avMMN contrast.

**EROS results**

Statistical maps of the EROS results are shown in Figure 4.6. The peak Z-scores, critical Z, and locations of the significant peak optical responses are summarized in Table 4.1. In the mixed response contrast, significant increases in optical signals were found in the inferior frontal gyrus (IFG) from 179 ms to 230 ms, the middle temporal gyrus (MTG) from 332 ms to 383 ms, and the superior temporal gyrus (STG) from 383 ms to 409 ms. In the avMMN contrast, similar to the mixed response contrast, increases in optical signals were found in the middle temporal gyrus (MTG) from 332 ms to 383 ms; however, different from the mixed response contrast, an early optical response in the superior temporal gyrus (STG) from 204 ms to 230 ms was found and there was no significant optical response in the frontal ROI.

Based on the location and latency shown in Table 4.1, the optical responses were grouped into four clusters: IFG(179-230), STG(204-230 and 383-409), MTG(332-383), and OCC(332-383). The peak amplitude of each cluster and its corresponding standard error are shown in Figure 4.7. Repeated measures ANOVA on the peak optical response with contrast and cluster as factors showed non-significant main effects of contrast ($F(1,12) = 2.556$, $p > .05$) and cluster ($F(4,48) = 0.742$, $p > .05$ with Greenhouse-Geisser correction, $\varepsilon = .481$), but a significant interaction effect ($F(4,48) = 4.126$, $p < .05$ with Greenhouse-Geisser correction, $\varepsilon = .527$), suggesting differences in the spatiotemporal dynamics of the brain responses between the two contrasts.

Paired t-tests (one tailed) were conducted to investigate if the optical responses in the mixed response contrast were larger than that in the avMMN contrast for the IFG(179-230), MTG(332-383), and OCC(332-383) clusters. A significantly larger optical response was found in

the IFG cluster (179-230)($t(15) = 1.799$, $p < .05$) and a marginally larger optical response was observed in the OCC cluster (332-383) ($t(14) = 1.709$, $p = .054$) for the mixed response contrast compared to the avMMN contrast. However, no significant difference in the optical responses between the contrasts was revealed in the MTG cluster (332-383) ($t(13) = -.103$, $p > .05$).

Due to the similarity of the location of the STG clusters (204-230and 383-409), repeated measures ANOVA with time of activation and contrast as factors was conducted to investigate the presence of a possible delay in the STG response. The main effects of time of activation $F(1,15) = 0.866$, $p > .05$) and contrast ($F(1,15) = 1.801$, $p > .05$) were non-significant but the interaction effect was ($F(1,15) = 6.171$, $p < .05$). Paired sample t-tests showed a larger optical response in the avMMN contrast from 204 ms to 230 ms ($t(15) = -2.292$, $p < .05$) and a larger optical response in the mixed response contrast from 383 ms to 409 ms ($t(15) = 1.887$, $p < .05$). This result suggested a delay in the STG response in the mixed response contrast.

**Discussion**

In the current study we adopted a passive oddball paradigm with a set of control conditions to dissociate AV integration and AV deviance detection. It was predicted that brain responses related to AV deviance detection would be captured by both the mixed response and the avMMN contrasts, while pure AV integration effects would be captured by the mixed response contrast. Furthermore, interaction between AV integration and avMMN would be reflected as a modulation of the amplitude or latency by different contrasts. Unique optical responses indicating AV integration were found in the IFG from 179 ms to 230 ms and the occipital cortex from 332 ms to 383 ms. Brain responses associated with AV deviance detection

were observed in the MTC from 332ms to 383 ms in both contrasts. Although there was no difference in the peak amplitude of the MTG response between the contrasts, the spatial and temporal extension of the MTG activity suggested possible interaction of the AV integration and AV deviance detection processes in the MTG. An interaction effect was found in the STG: instead of modulating the amplitude of the brain response, the latency of the STG activity was different between the contrasts. The STG activity occurred earlier for the avMMN contrast (from 204ms to 230ms) than for the mixed response contrast (from 383 ms to 409ms).

The EROS results demonstrated a frontal-temporal/occipital activation pattern in AV integration. Previous ERP and fMRI studies (e.g., Saint-Amour, De Sanctis, Molholm, Ritter, & Foxe, 2007; Skipper, van Wassenhove, Nusbaum, & Small, 2007) suggested a similar frontal-temporal activation pattern in AV speech perception. However, the current study offered two advantages. First, the passive oddball paradigm allowed disambiguation of the AV integration from the AV deviance detection process. Second, the fast optical signal measured revealed spatiotemporal dynamics without having to infer them from sluggish hemodynamic signals. The sequential activation pattern, frontal followed by temporal and occipital cortex activities, suggested top-down control of the primary or association sensory cortices by the frontal cortex or a feedback mechanism from the frontal to the sensory cortex. More importantly, we demonstrated the interaction of AV integration and AV deviance detection processes. In other words, the additive factor model assumption used to analyze different cognitive processes may not be valid. This kind of interaction effect is important for understanding differences in experimental results measured with different tasks or experimental designs in AV integration or speech perception (see Calvert, 2001 for review).

Frontal cortex involvement in AV integration was found in fMRI studies (Callan, et al., 2003; Miller & D'Esposito, 2005; Skipper, Nusbaum, & Small, 2005; Skipper, van Wassenhove, Nusbaum, & Small, 2007) and is thought to index the difficulty in AV integration. For example, IFC activity was found to increase with the difficulty to formulate a fused percept of AV incongruent speech (Miller & D'Esposito, 2005; Skipper, van Wassenhove, Nusbaum, & Small, 2007). On the other hand, the frontal activity can be suppressed if one can identify the syllable easily by the auditory channel alone (Skipper, van Wassenhove, Nusbaum, & Small, 2007). Further analysis suggests that the IFC activity in the mixed response contrast is contributed by a decrease of frontal response to the AV standards (Figure 4.8). This indicates suppression in the AV integration process with repeated AV speech stimuli. A similar suppression effect was found in the STC from 383ms to 409ms in the congruent AV standards (Figure 4.8), consistent with results from ERP studies (Besle, Fort, Delpuech, & Giard, 2004; van Wassenhove, Grant, & Poeppel, 2005).

Although IFC activation is commonly found in speech perception, its functional role remains controversial. Based on the motor theory of speech perception and the analysis-by-synthesis model , Skipper et al. (Skipper, van Wassenhove, Nusbaum, & Small, 2007) proposed that the IFC, as part of the speech production system, is required for speech perception. Specifically, information generated from the speech production system was used to constrain the percept of AV speech. Further support for the motor theory was provided by frontal and primary motor cortex activities in response to speech sound (see Fadiga, Craighero, & Olivier, 2005 for review). However, other studies (e.g., Sundara, Namasivayam, & Chen, 2001) found frontal and motor cortex activation in observing speech movement, but not in listening to speech sound. Another possible functional role of the IFC activation is that it might reflect response selection as

suggested in a TMS study (Tremblay & Gracco, 2009). The left IFC may be recruited for top down control to resolve incongruent linguistic stimuli (January, Trueswell, & Thompson-Schill, 2009). Specifically, the IFC is responsible for selecting among competing representations (selection or interference resolution) (Nelson, Reuter-Lorenz, Persson, Sylvester, & Jonides, 2009). Although the frontal-followed-by-sensory cortex activation pattern is consistent with the prediction in Skipper et al. (2007), the experimental design of the current study does not allow us to conclude on the functional role of the frontal cortex in AV integration.

The ERP results are consistent with the prediction that a more sustained and larger brain response is involved in the mixed response contrast compared to the avMMN contrast, as both AV integration and AV deviance detection are captured by the mixed response contrast. The ERP results are also consistent with previous ERP and MEG studies that used the passive oddball paradigm (e.g., Colin, et al., 2002; Mottonen, Krause, Tiippana, & Sams, 2002; Saint-Amour, De Sanctis, Molholm, Ritter, & Foxe, 2007; Sams, et al., 1991); both the latency and the amplitude of the negativity (i.e. > 1μV) recorded in the mixed response contrast were similar to those observed in previous studies that had used a similar contrast (Saint-Amour, De Sanctis, Molholm, Ritter, & Foxe, 2007). In the avMMN contrast of the current study, incongruent AV deviants were compared to incongruent McGurk stimuli in the AV control block, such that the deviance detection response could be measured without being contaminated by a congruency effect. We found a negativity in the avMMN contrast similar in size (< 1μV) to that reported in Winkler et al. (2009) when congruency was also controlled, but smaller than that reported in Saint-Amour et al. (2007) when congruency was not controlled.

The most prominent difference between the ERP aMMN and avMMN is the absence of mastoid inversion with the nose reference in the avMMN. Colin et al. (2002) collected data from

a single subject in an extended recording session to show that the absence of mastoid inversion in avMMN was not related to a poor signal-to-noise ratio or a lack of statistical power. Their results implied a difference in the location or orientation of the dipoles in the aMMN and avMMN responses, although mastoid inversion is not a necessary feature of aMMN. Due to the absence of a condition consisting of only auditory speech stimuli in the current study, we cannot directly compare brain responses of aMMN and avMMN. However, the left STC avMMN activity found in the current study is consistent with previous ERP and MEG aMMN studies that used language stimuli (Näätänen, et al., 1997; Pulvermuller & Shtyrov, 2006; Shtyrov & Pulvermuller, 2007). Studies comparing MMN elicited by acoustic and phonetic stimuli also suggested right lateralization for acoustic stimuli and left lateralization for phonetic stimuli (Tervaniemi, et al., 2000).

Compared to previous optical aMMN studies (Tse & Penney, 2007; Tse, Tien, & Penney, 2006), the STC and MTC optical responses observed in the current study are relatively posterior and inferior to the left homologous position of the right lateralized optical aMMN, respectively. Such results suggest that higher level of processing in the secondary auditory cortex or association areas is needed in AV deviance detection. A similar hierarchical model was suggested by Fuster et al. (2001; 2006; 2000).

In Fuster's model, the central fissure divides the brain into posterior perceptual and anterior executive systems. Each system is hierarchically organized. Sensory information transfers from the primary sensory cortex to the association areas (temporal and parietal cortex) to formulate more complex representations by integrating basic sensory percepts. In other words, the association areas take input from multiple sensory areas to form multisensory percepts. Feedback mechanisms from the association areas to the sensory cortex are also present, allowing

top down modulation or prediction of the sensory percepts. The difference in the location of the optical aMMN and avMMN responses in the temporal cortex suggested the involvement of multisensory areas in AV deviance detection. Furthermore, the prefrontal cortex in the executive system is responsible for cross integration of information across time and executive function, while the premotor and motor cortices are responsible for selecting motor responses and the execution of motor commands. Direct connection between the perceptual and executive systems allows top down modulation by the frontal cortex on the perceptual system i.e. the temporal cortex as observed in the EROS data. The presence of these networks was supported by animal studies that used anterograde tracers to demonstrate direct connections between the primary visual cortex and the auditory association cortex (Rockland & Ojima, 2003) and projections from the frontal cortex to auditory association cortex (Hackett, Stepniewska, & Kaas, 1999; Kaas & Hackett, 2000).

In summary, the current study showed that AV integration and AV deviance detection mechanisms can be dissociated with proper control conditions. A frontal-temporal/occipital activation pattern suggesting a top-down process or feedback system was observed in AV integration. More interestingly, AV integration and AV deviance detection may interact with each other, leading to modulation in the amplitude or latency of the brain responses.

# Chapter 5

## General Discussion and Conclusion

Previous ERP studies suggested that deviance detection can take place at various stages along the information processing stream and it involves very different cognitive processes. Three experiments were conducted to study the interactions between the frontal and temporal cortices in 1) processing semantic and syntactic violations, 2) representing regularities, and 3) detecting audiovisual deviance. Event-related Optical Signal (EROS) was measured in these experiments to reveal the spatiotemporal dynamics of the brain activities during different kinds of deviance detection processes.

The first experiment investigated brain responses to semantic and syntactic violations in language processing. By using EROS, the rapid interactions between the left superior/middle temporal cortices (S/MTC) and inferior frontal cortices (IFC) during the processing of semantically or syntactically anomalous sentences were revealed. Simultaneous ERP recordings showed activities that are typically elicited when processing anomalous stimuli: the N400 and the P600 for semantic and syntactic anomalies, respectively. The EROS elicited by semantically anomalous words revealed an increase in S/MTC activity (corresponding in time to the N400), followed by IFC activity. Syntactically anomalous words evoked a similar brain activation sequence, with S/MTC activity (corresponding in time to the P600) followed by frontal activity. However, the S/MTC activity corresponding to detecting semantic anomalies was more ventral than that corresponding to detecting syntactic anomalies. These data suggest that activations related to processing anomalies in sentences proceed from temporal to frontal brain regions for both semantic and syntactic anomalies. More prominent re-entry activity from the frontal to

temporal cortex was observed with semantic anomalies, suggesting possible feedback mechanisms between the frontal and temporal cortices. Being the first study that used EROS to investigate language processing, it successfully demonstrated that EROS can be used to image rapid interactions across cortical areas.

The second experiment shifted the focus from deviance detection to regularity representation. Specifically, EROS was used to differentiate brain responses, in both time and space, related to implicit rules with different stimulus contingencies and processing requirements. Three contingency rules, Global Probability, Single Repetition, and Local Probability, which differed in two dimensions, template mismatch and memory updating, governed the sequence of high and low frequency auditory tones. Planned comparisons with orthogonal contrasts between different trial types revealed temporal activities at 200-300 ms. Such activities reflect a template mismatch response for the Global Probability and Single Repetition rules. At later intervals (300-500 ms), a fronto-parietal network was found to be sensitive to the Global Probability and Local Probability rules which required memory updating. Event-related brain potentials (ERP) data, recorded in a separate session, were consistent with the EROS results. This study suggests that the brain can simultaneously hold different models of stimulus contingencies within the information processing stream, but these representations are held at different levels, both in terms of the latency and location of the brain responses, according to the processing requirements of the contingency rules.

The third experiment extended the investigation of brain responses in deviance detection from unimodal to multimodal sensory systems. Deviance detection has been widely used as a tool to study various cognitive processes; however, it is difficult to separate it from the cognitive processes being studied. Another purpose of this study was to dissociate audiovisual deviance

detection from audiovisual integration in speech perception with a set of control conditions. The results indicate that audiovisual deviance elicits a short duration response in the middle/superior temporal gyrus, whereas conflict/integration manipulations elicit a more extended response involving also the inferior frontal and occipital regions. More interestingly, the interaction of audiovisual deviance detection and audiovisual integration processes was revealed by temporal modulation of the temporal activity.

Consistent with previous studies (see Fabiani, 2006 for review), the current study showed that deviance detection can take place at various stages of cognitive processing. Deviance detection is not a unitary process, but a collection of processes which operate at different levels in cognition. In other words, deviance detection is a general property of all cognitive processes. As cognitive processes are very different from each other, it is not surprising that different brain regions or neural substrates are involved in different kinds of deviance detection. The exact location of brain regions involved may vary, however, the temporal and frontal cortices are consistently activated in various kinds of deviance detection processes as shown in the current thesis and previous EROS studies (Low, Leaver, Kramer, Fabiani, & Gratton, 2006; Sable, et al., 2007; Tse & Penney, 2007; Tse & Penney, 2008; Tse, Tien, & Penney, 2006). The frontal and temporal cortices are the basic neural substrates, a general deviance detection system, that are shared among various cognitive processes.

The hierarchical organization model of the brain suggested by Fuster et al. (2001; 2006; 2000) provides a framework for inferring the functional roles and properties of these basic processing units in the temporal and frontal cortices. In the posterior sensory system, sensory information is transferred from the primary sensory cortex to the association areas (temporal and parietal cortex) to formulate more complex representations by integrating basic sensory percepts.

The temporal cortex activity commonly found in deviance detection suggested that deviance detection can occur at the sensory level, probably as a result of comparisons between the sensory percepts of the standard and the deviant. The location at which the comparison actually takes place depends on the complexity of the percepts. For example, deviants in physical properties (frequency or intensity) of a tone are detected in the primary or secondary auditory cortex, while semantic and syntactic violations in sentences, or deviants in the audiovisual percepts of speech sound, are perceived in higher level association areas (MTC or posterior STC). In short, the deviance detection mechanism compares sensory percepts at different levels along the processing stream in the posterior sensory system.

In Fuster's model, the anterior executive system is responsible for executive function and the integration of information across time. More importantly, the anterior executive system interacts with the posterior sensory system to manipulate internal representations (sensory percepts and contents of the long term memory store). An intriguing possibility is that this process may be a critical component of working memory. Similarly, we can conceptualize the frontal activity in deviance detection as a reflection of manipulations of internal representations, ranging from updating or fine-tuning the template for deviance detection to remolding the deviant for comparisons based on past experience. This idea is consistent with various studies suggesting that the frontal cortex, specifically the IFC, is involved in conflict resolution (January, Trueswell, & Thompson-Schill, 2009; Nelson, Reuter-Lorenz, Persson, Sylvester, & Jonides, 2009) or contrast enhancement (Opitz, Rinne, Mecklinger, Cramon, & Schröger, 2002). The kind of manipulations implemented may lead to variations in the locations of the brain responses.

The current thesis used EROS as the major brain imaging technique to demonstrate the importance of uncovering the spatiotemporal dynamics in various cognitive processes. The

difference in activation sequence among the various deviance detection processes reflects differences in how the basic cognitive components in frontal and temporal cortices interact with each other. Frontal-temporal or temporal-frontal activation patterns were commonly found in deviance detection. More interestingly, re-entry patterns, frontal-temporal-temporal or temporal-frontal-temporal activation sequences were also observed in various EROS studies. The re-entry pattern may reflect a refitting or reanalyzing mechanism that takes place when internal representations are passed between the frontal and temporal cortices for cyclical remolding and comparison. However, it is also possible that the frontal and temporal cortices represent two independent systems that work in parallel. Having multiple deviance detection mechanisms increases the likelihood of a deviant stimulus being picked up by at least one of them.

In summary, the current thesis applied EROS to image the interactions between the frontal and temporal cortices in 1) detecting semantic and syntactic violations, 2) representing regularities, and 3) detecting audiovisual deviance. The results suggested that although deviance detection takes place at various cognitive stages, similar basic processing units in the frontal and temporal cortices are shared among them.

## Tables

Table 2.1. Statistically significant peak EROS responses

| ROI time, msec | *x, y, z** | Peak Z (Z critical) | Location† | BA† |
|---|---|---|---|---|
| Semantic | | | | |
| S/MTC-179 | –50, –50, –1 | 2.92 (2.67) | ITG | 37 |
| S/MTC-204 | –50, –47, –3 | 3.15 (2.72) | Subgyral | 37 |
| S/MTC-384 | –52, –25, –3 | 2.98 (2.84) | STG | 21 |
| IFC-512 | –47, 4, –10 | 2.68 (2.68) | STG | 38 |
| IFC-563 | –65,14, 18 | 2.99 (2.92) | IFG | 45 |
| IFC-588 | –65, 24, 16 | 2.96 (2.88) | IFG | 45 |
| S/MTC-665 | –72, –25, 22 | 3.11 (2.89) | STG | 42 |
| IFC-717 | –47, 22, 1 | 3.10 (2.91) | IFG | — |
| FC-742 | –62, 27, –6 | 3.46 (2.97) | IFG | 47 |
| IFC-768 | –60, 24, –3 | 3.14 (2.02) | IFG | 47 |
| S/MTC-844 | –72, –22, 22 | 2.75 (2.73) | IPL | 40 |
| S/MTC-870 | –72, –22, 22 | 2.96 (2.78) | IPL | 40 |
| IFC-844 | –65, 4, 4 | 3.13 (2.87) | PCG | 6 |
| IFC-870 | –65, 4, 4 | 3.10 (2.90) | PCG | 6 |
| S/MTC-896 | –50, –52, –1 | 3.07 (2.85) | ITG | 37 |
| IFC-972 | –50, 37, 13 | 3.08 (2.96) | MFG | 46 |
| IFC-998 | –52, 37, 11 | 3.00 (2.94) | IFG | 46 |
| Syntactic | | | | |
| S/MTC-256 | –67, –52, 22 | 3.00 (2.73) | SMG | 40 |
| S/MTC-563 | –65, –55, 22 | 2.99 (2.81) | STG | 22 |
| S/MTC-588 | –60, –45, 19 | 2.80 (2.77) | STG | 22 |
| S/MTC-819 | –52, –22, 7 | 3.36 (2.95) | STG | 41 |
| S/MTC-844 | –50, –20, 9 | 3.42 (2.87) | TTG | 41 |
| S/MTC-870 | –50, –20, 19 | 3.05 (2.80) | Insula | 13 |
| S/MTC-947 | –47, –22, 6 | 2.93 (2.73) | STG | 41 |
| S/MTC-972 | –62, –25, –1 | 3.12 (2.86) | MTG | 21 |
| IFC-972 | –55, 34, 8 | 3.04 (3.04) | IFG | 46 |
| IFC-998 | –55, 34, 8 | 3.11 (3.04) | IFG | 46 |

ROI, region of interest; BA, Brodmann's area; ITG, inferior temporal gyrus; STG, superior temporal gyrus; IFG, inferior frontal gyrus; IPL, inferior parietal lobe; PCG, precentral gyrus; MFG, middle frontal gyrus; MTG, middle temporal gyrus; SMG, supramarginal gyrus; TTG, transverse temporal gyrus. Shaded/clear rows represent statistically significant EROS responses belonging to the same cluster; nonsignificant responses are not included. Shading of the same color allows visual clustering of consecutive significant time points within the same region.

**x, y, z* are in Talairach coordinates.

†Brodmann's areas and corresponding locations of the brain are obtained from the Talairach Daemon, which shows the nearest grey matter to the peak EROS response.

Table 2.2. Stepwise multiple regression analyses showing a double-dissociation pattern in the relationship between EROS and ERP effects

| DV*/ERP | Independent variables† (predictors) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Semantic EROS | | | | Syntactic EROS | | | |
| N400 | Predictors | β‡ | *t*(12) | *P* | Significant predictors not found<br>$F(2,13) = 1.66$, ns§ | | | |
| | S/MTC-179 msec | 0.76 | 7.07 | <0.001 | | | | |
| | S/MTC-384 msec | –2.27 | –4.77 | <0.005 | | | | |
| | IFC-512ms | 1.37 | 3.14 | <0.05 | | | | |
| | $R^2 = 0.884$, adj. $R^2 = 0.835$<br>$F(3,7) = 17.88$, $P < 0.001$ | | | | | | | |
| P600 | Significant predictors not found<br>$F(3,7) = 1.05$, ns¶ | | | | Predictors | β‖ | *t*(12) | *P* |
| | | | | | S/MTC-819 msec | 0.46 | 3.83 | <0.005 |
| | | | | | S/MTC-972 msec | –0.52 | –2.48 | <0.05 |
| | | | | | $R^2 = 0.608$, adj. $R^2 = 0.543$<br>$F(2,12) = 9.31$, $P < 0.005$ | | | |

*Peak ERP amplitude at Pz electrode for each subject; search windows were 200–500 ms for N400 and 500-1500 ms for P600.

†Averaged optical signal of a 0.5 cm × 0.5 cm × 0.625 cm ROI around the peak *Z* score.

‡A negative $\beta$ is expected in predicting N400 because of its negative polarity; the significant predictor consistent with this expectation is highlighted.

§Using predictors S/MTC-819 ms, S/MTC-972 ms of Syntactic EROS to predict N400; ns, not significant.

¶Using predictors S/MTC-179 ms, S/MTC-384 ms, IFC-512 ms of Semantic EROS to predict P600; ns, not significant.

‖A positive $\beta$ is expected in predicting P600 because of its positive polarity; the significant predictor consistent with this expectation is highlighted.

Table 3.1.

Behavioral results (reaction time) - contrast analysis

| | Global Probability | | Single Repetition | | Local Probability | |
|---|---|---|---|---|---|---|
| Sessions | Sum(SD) | *t* | Sum(SD) | *t* | Sum(SD) | *t* |
| 1st ERPs | 34.10(80.81) | 1.69 | 1.98(44.79) | 0.18 | 0.25(17.61) | 0.06 |
| 2nd ERPs | 125.46(78.53) | 6.39* | 33.14(50.92) | 2.60* | 21.22(38.62) | 2.29* |
| Combined EROS | 167.20(93.90) | 7.12* | 25.01(29.76) | 3.36* | 21.60(6.04) | 5.39* |

Note: * $p < .05$ (1-tailed), $df$=15

Table 3.2.

ERPs results - contrast analysis with 200-250ms and 250-450ms time windows

| Time Windows | Global Probability | | Single Repetition | | Local Probability | |
|---|---|---|---|---|---|---|
| N2 (200-250ms) | Sum(SD) | *t* | Sum(SD) | *t* | Sum(SD) | *t* |
| Fz | -6.43(6.00) | -4.29* | -1.10(3.72) | -1.19 | -0.27(1.64) | -0.66 |
| Cz | -8.97(7.66) | -4.69* | -1.95(4.33) | -1.81* | -0.11(2.16) | -0.21 |
| Pz | -4.67(5.66) | -3.30* | -1.20(3.60) | -1.34 | -0.23(2.07) | -0.44 |
| P3 (250-450ms) | Sum(SD) | *t* | Sum(SD) | *t* | Sum(SD) | *t* |
| Fz | 4.19(6.22) | 2.69* | -0.08(3.00) | -0.11 | 0.21(1.68) | 0.49 |
| Cz | 0.63(3.42) | 7.19* | 1.37(2.40) | 0.73 | 1.37(2.40) | 2.28* |
| Pz | 9.91(5.09) | 7.78* | 1.21(3.04) | 1.59 | 1.57(2.15) | 2.92* |

Note: * $p < .05$ (1-tailed), $df$=15; the t-test for N2 is negative because N2 is a negative component.

Table 4.1. Statistically significant peak EROS responses

| | AV Deviant versus AV Standard Interaction Contrast | | | | | AV Deviant versus AV Control Contrast | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| time(ms) | ROI | Coordinate | Peak Z (critical Z) | Location | BA | ROI | Coordinate | Peak Z (critical Z) | Location | BA |
| 179-204 | Frontal | 38, -1 | 3.16 (2.40) | IFG | 45,46 | - | - | - | - | - |
| 204-230 | Frontal | 38, -1 | 2.87 (2.60) | IFG | 45,46 | Temporal | -51,14 | 2.70 (2.50) | STG | 22 |
| 332-358 | Temporal | -33, -8 | 2.94 (2.43) | MTG | 21 | Temporal | -26,-3 | 2.50 (2.43) | MTG | 21 |
| | Occipital | -13, 4 | 2.84 (2.59) | Cuneus | 17,18 | | | | | |
| 358-383 | Temporal | -33, -8 | 2.96 (2.37) | MTG | 21 | - | - | - | - | - |
| | Occipital | 2, 9 | 2.83 (2.70) | Cuneus | 17,18 | | | | | |
| 383-409 | Temporal | -38, 17 | 2.56 (2.30) | STG | 22 | - | - | - | - | - |

Note: ROI, region of interest; BA, Brodmann's area; STG, superior temporal gyrus; IFG, inferior frontal gyrus; MTG, middle temporal gyrus. Brodmann's areas and corresponding locations of the brain are obtained from the Talairach Daemon, which shows the nearest grey matter to the peak EROS response.

**Figures**

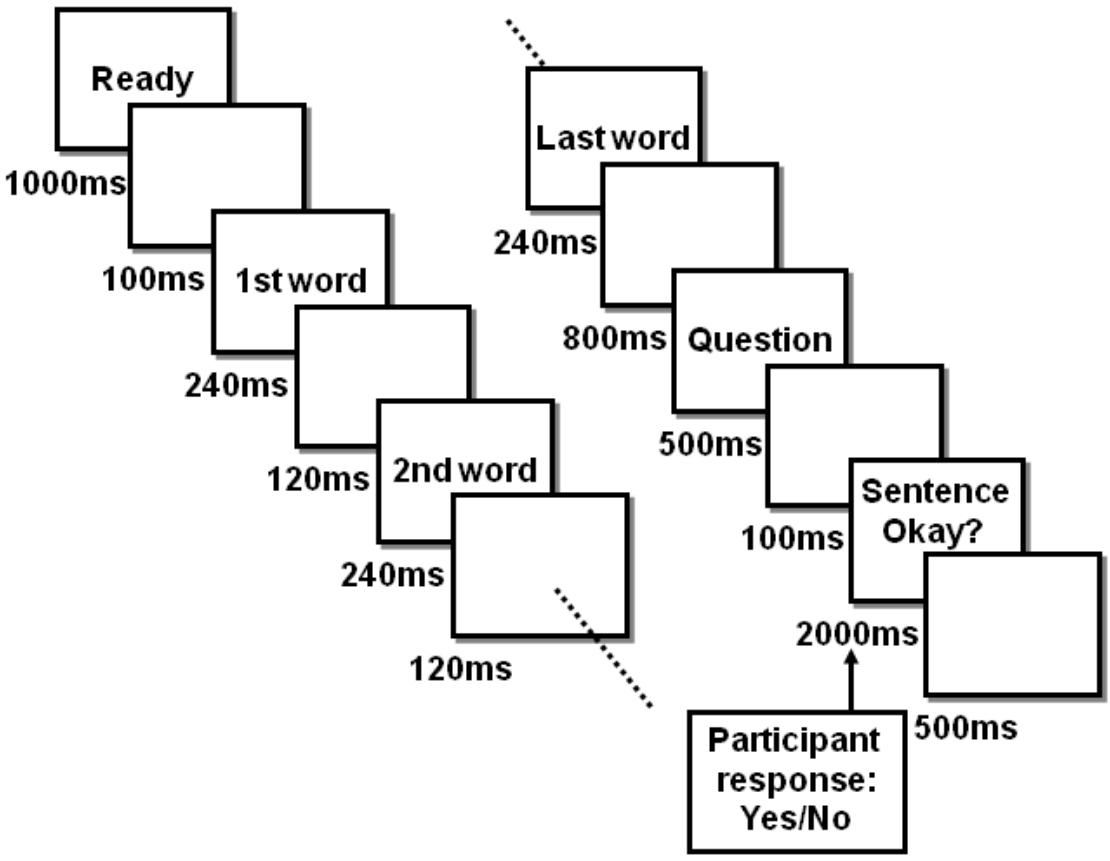


Figure 2.1. Sequential presentation of the stimuli making up each sentence and duration of each display.

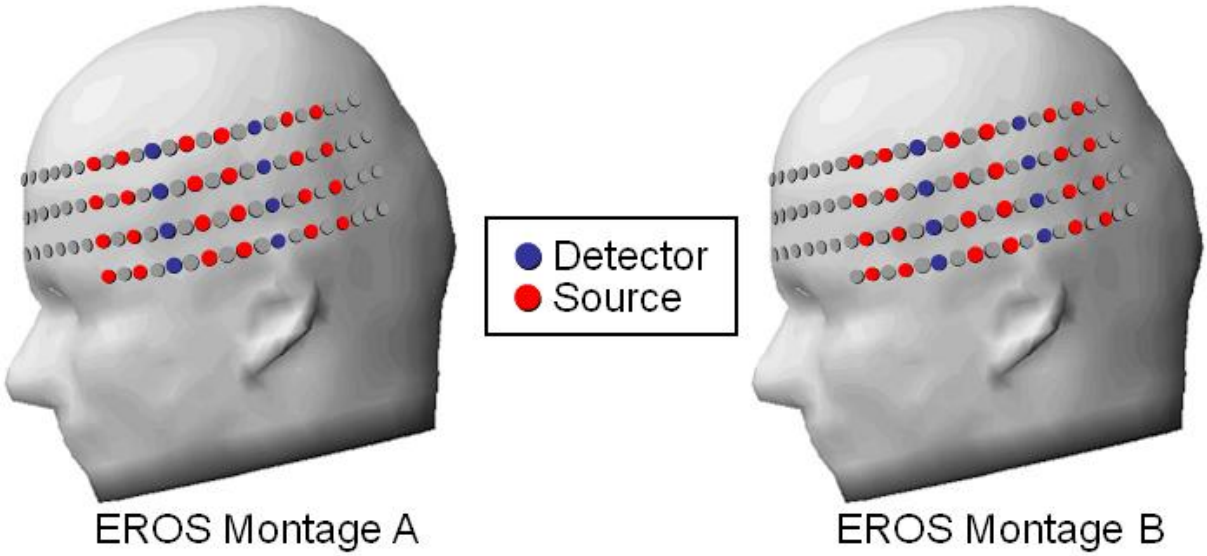


Figure 2.2. EROS recording montages. Montage A has the same pattern as Montage B, but one column (1.7cm) anterior to Montage B. This arrangement improves the spatial resolution of EROS recording.

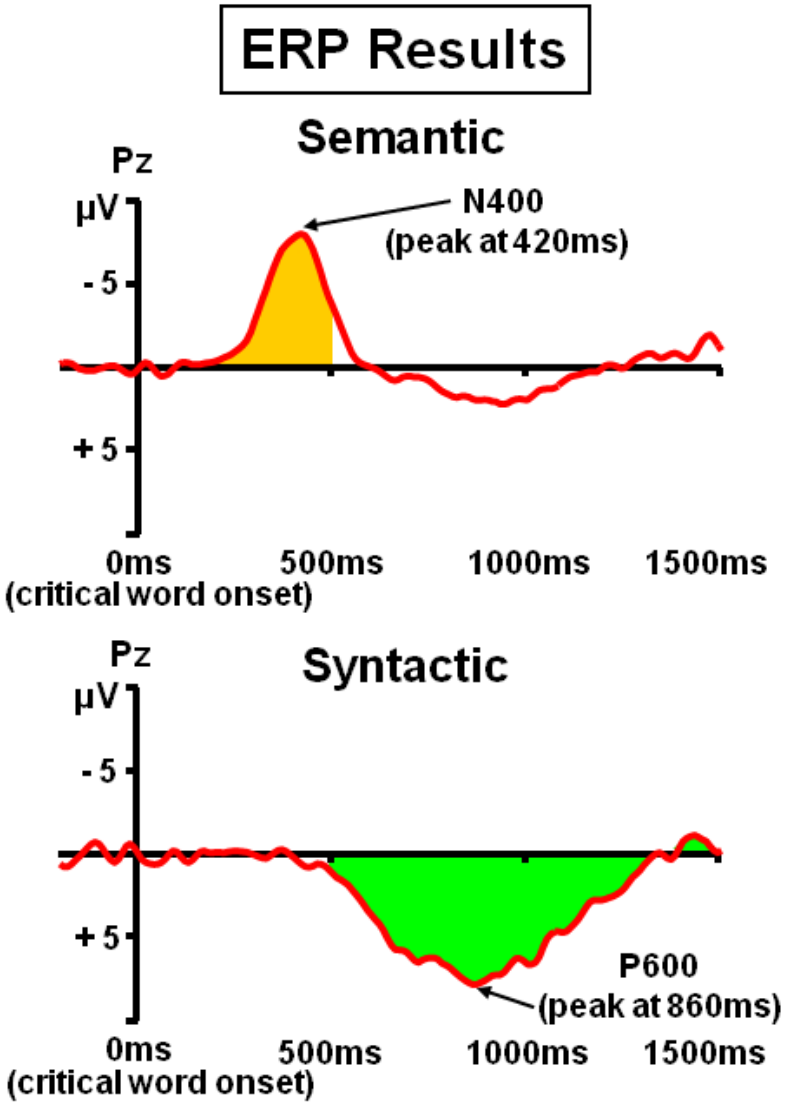


Figure 2.3. Grand average ERP difference waveforms (unacceptable minus acceptable) at the Pz electrode, time-locked to critical word onset. (*Upper*) For the semantic condition, the N400 effect is shown in orange. (*Lower*) For the syntactic condition, the P600 effect is shown in green.

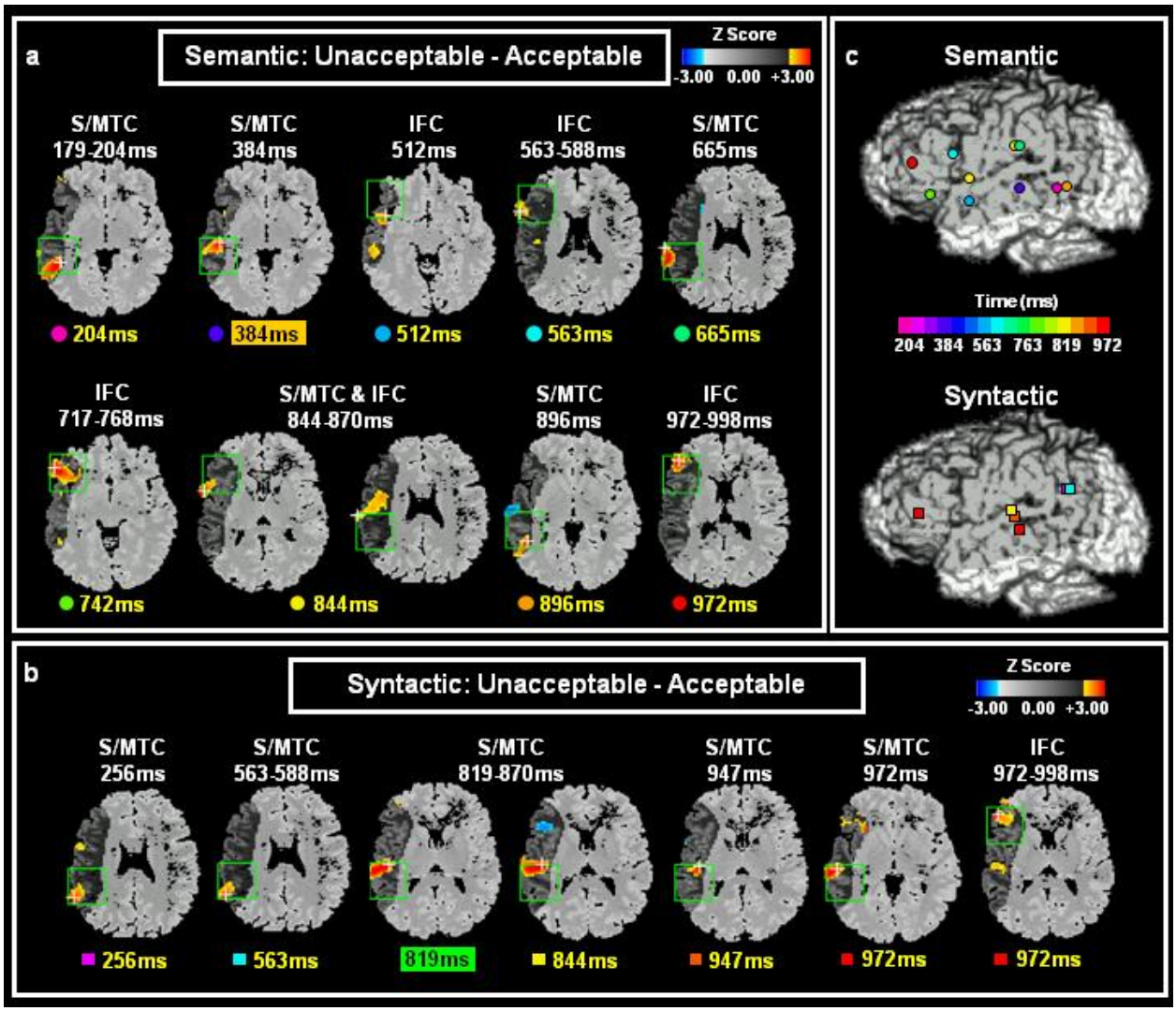


Figure 2.4. Statistical maps of EROS data. (*a* and *b*) Axial slices with significant increases in optical signal for semantic and syntactic anomalies, respectively. Green boxes indicate the locations of the 4-cm cubes defining each ROI. The white cross within the ROI indicates the location of peak activity. The dark gray shading shows the area of cortex interrogated by the montage. Significant EROS activities are typically found at similar locations across more than five axial slices (1.25-mm thickness for each slice) and across adjacent time windows (25.6 ms). EROS activities at similar locations and time windows were regarded as temporospatial clusters (white labels above each slice). Only the slices with the most significant peak activity for each temporospatial cluster are shown here, except for the 819-ms syntactic response (whose cluster peaks at 844 ms). The latencies of the peak optical activities are indicated in yellow below each slice. S/MTC 384 ms in the semantic condition and S/MTC 819 ms in the syntactic condition (corresponding to the N400 and P600 ERP effects, respectively) are highlighted. (*c*) Relative positions of the peak activities of the temporospatial clusters plotted over left brain views. The semantic and syntactic conditions are shown in *Upper* and *Lower*, respectively. The latency of the EROS response is color coded.

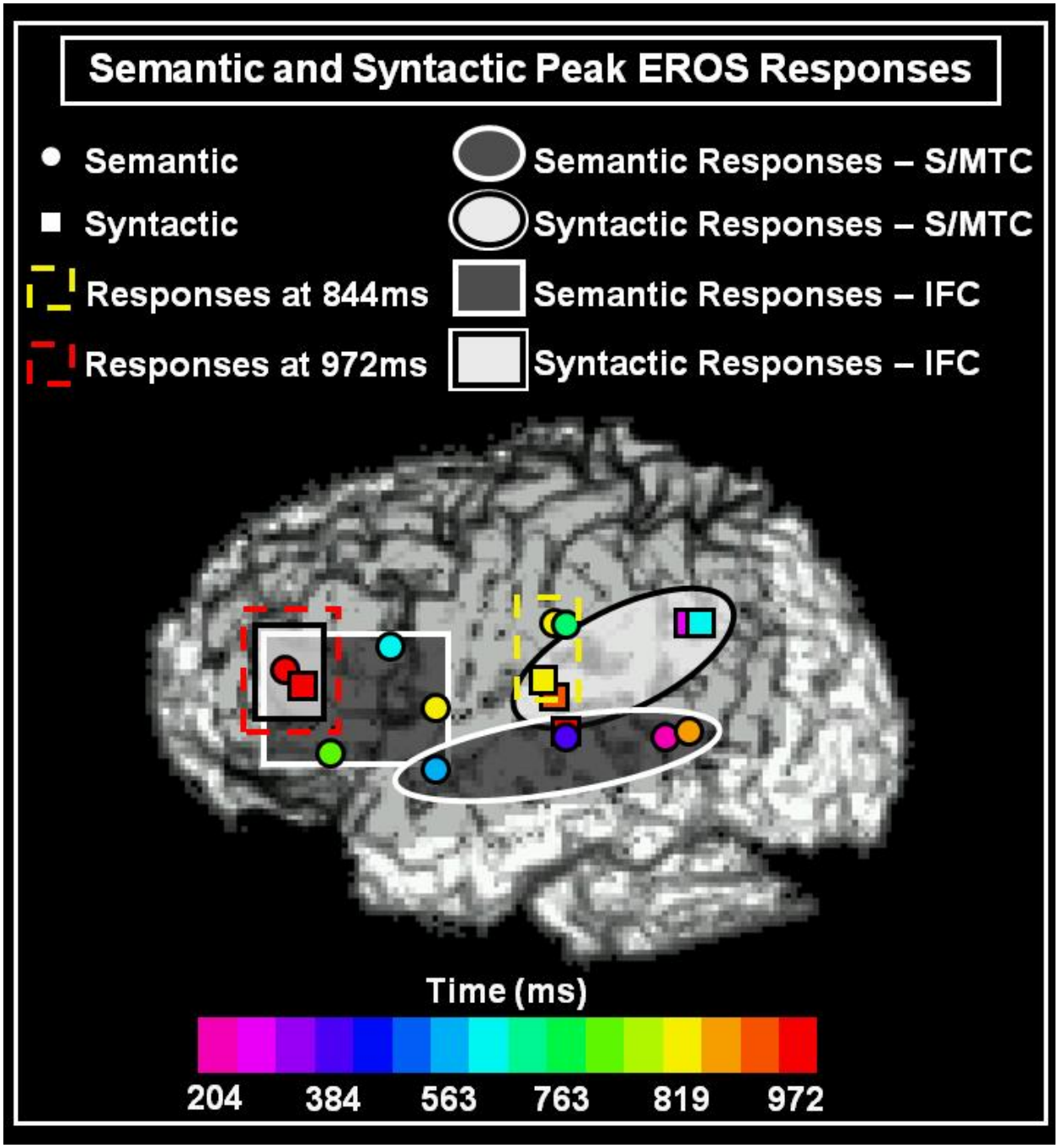


Figure 2.5. Semantic (small circles) and syntactic (small squares) peak responses for each temporospatial cluster overlaid on the same template brain. The latency of the response is color-coded. Regions activated by semantic anomalies are dark-gray-shaded; those activated by syntactic anomalies are light gray-shaded (both ellipses, temporal; rectangle, frontal). Convergence regions are indicated by the dashed rectangles.

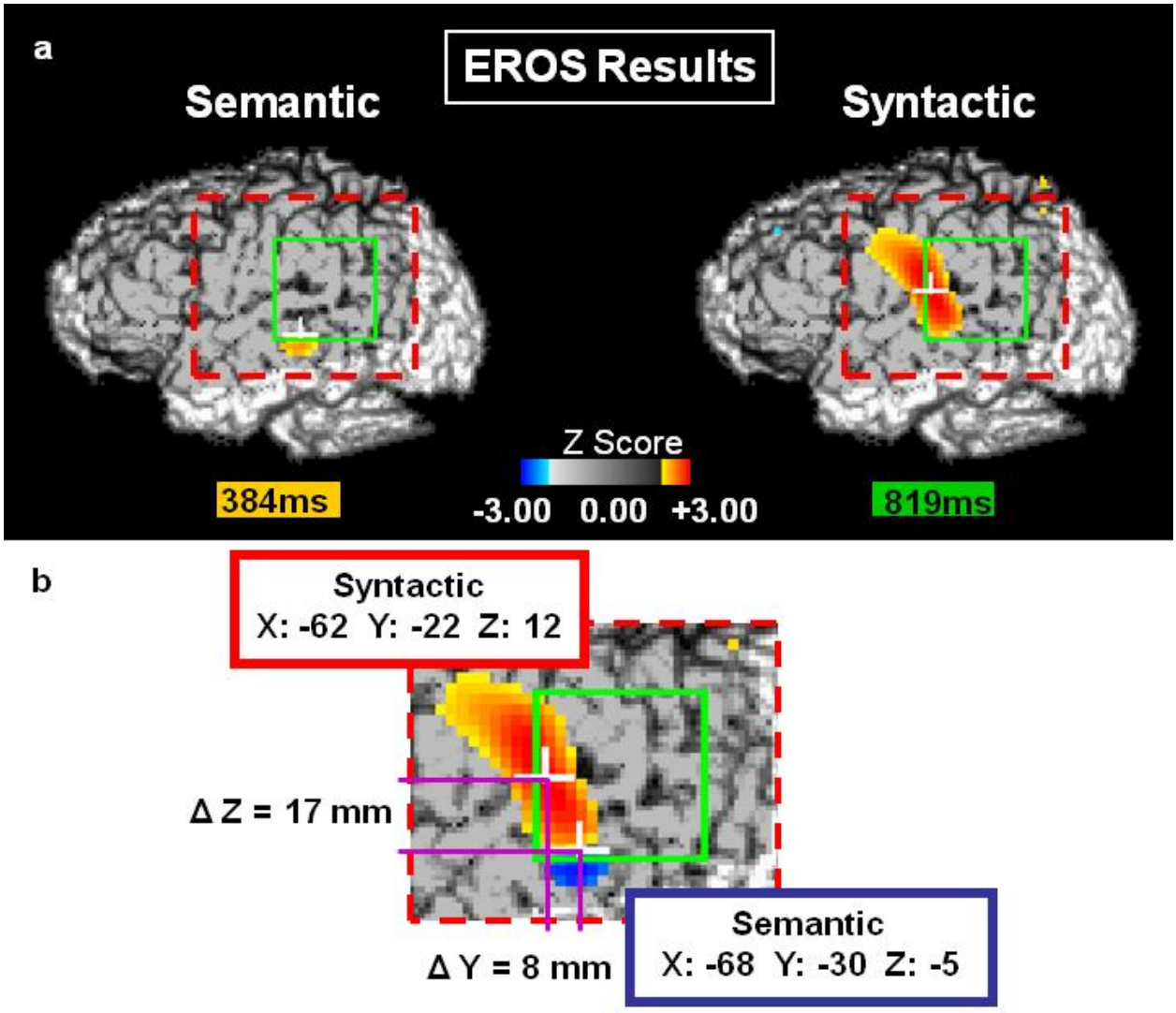


Figure 2.6. Test of location of ERP-predicting EROS responses. (*a*) Statistical maps of EROS results corresponding to the N400 and P600 ERP effects. The green boxes indicate the ROI for the STC, and the red boxes indicate the region shown in *b*. (*b*) Enlargement of the region within the red box in *a*, in which the areas showing the effects for semantic and syntactic conditions are overlaid. Blue, EROS activity in the semantic condition; red, EROS activity in the syntactic condition.

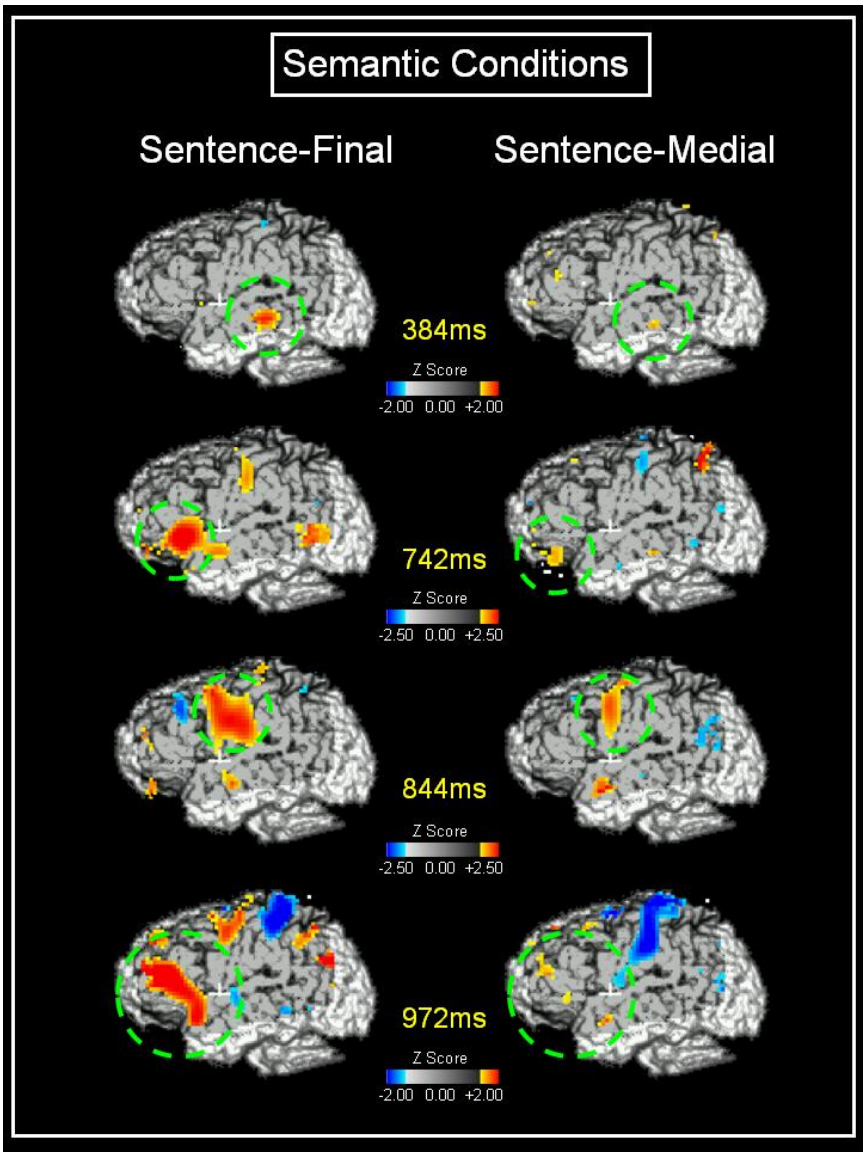


Figure 2.7. Comparison of surface-projected brain activity elicited by semantic deviance in the middle (right column) and at the end (left column) of a sentence.

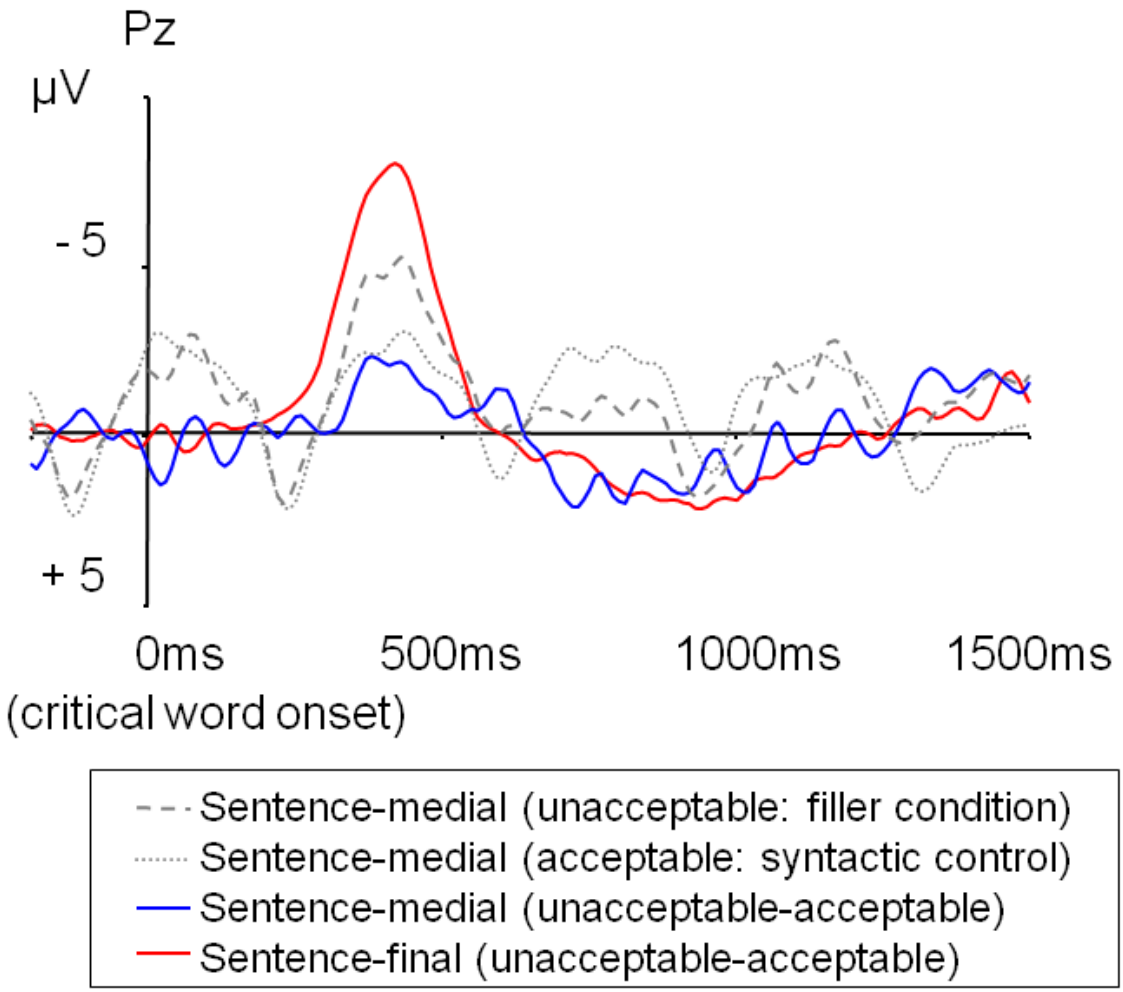


Figure 2.8. ERP waveforms at Pz for different types of words occurring in the middle or at the end of a sentence (unacceptable = semantic anomalies). Red and blue solid lines designate subtraction waveforms; gray dotted and dashed lines are the original unsubtracted waveforms for the midsentence condition.

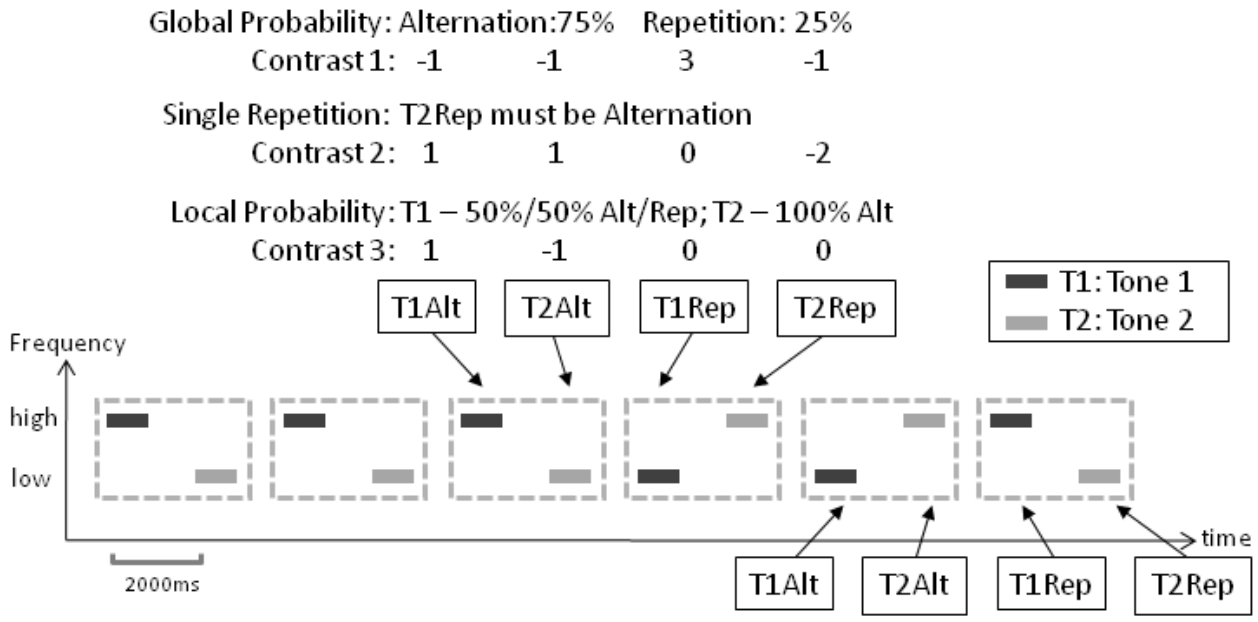


Figure 3.1. High and low frequency tones were presented to the participants and the sequence of the stimulus was governed by three contingency rules, global probability, single repetition, and local probability rules. Three orthogonal contrasts were setup for capturing activities related to the three contingency rules.

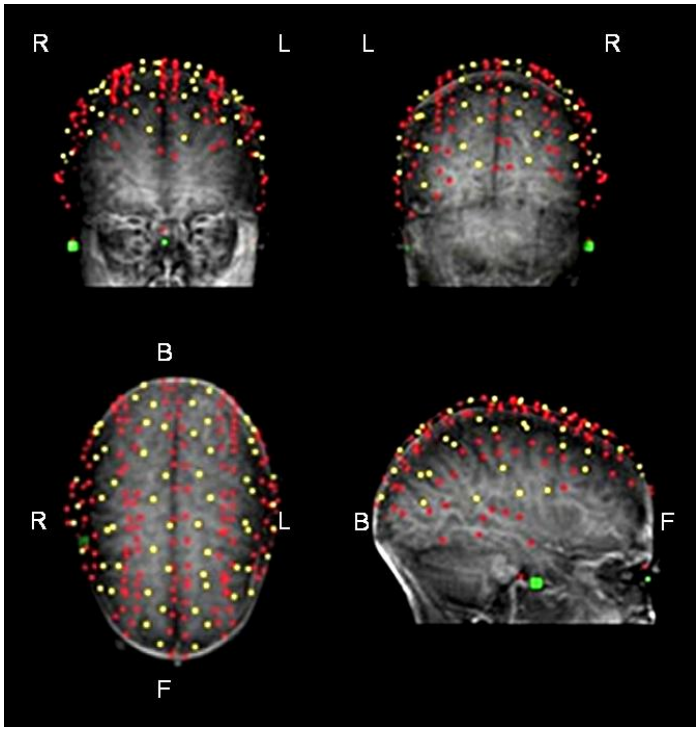


Figure 3.2. Anterior coronal (top-left), posterior coronal (top-right), superior axial (bottom-left), and right lateral (bottom-right) view of the EROS recording montages. The positions of the light source fibers (red dot), detector fibers (yellow dot), the nasion points (green dot) and pre-auricular points (green dot) were coregistered on the structural MRI for each participant. Four montages were used to record data from a total of 768 source-detector pairs. F:front; B:Back; L:Left; R:Right.

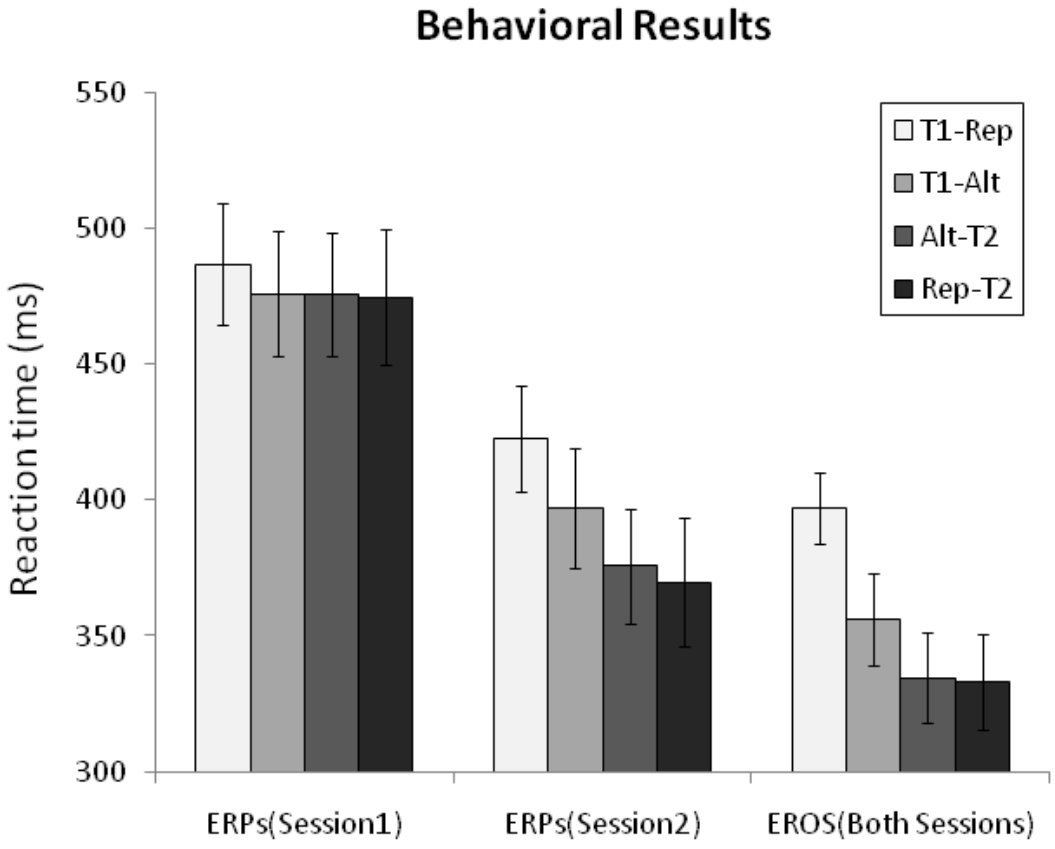


Figure 3.3. Average reaction times for each of the four trial types for the first, second ERP sessions, and the combined EROS session.

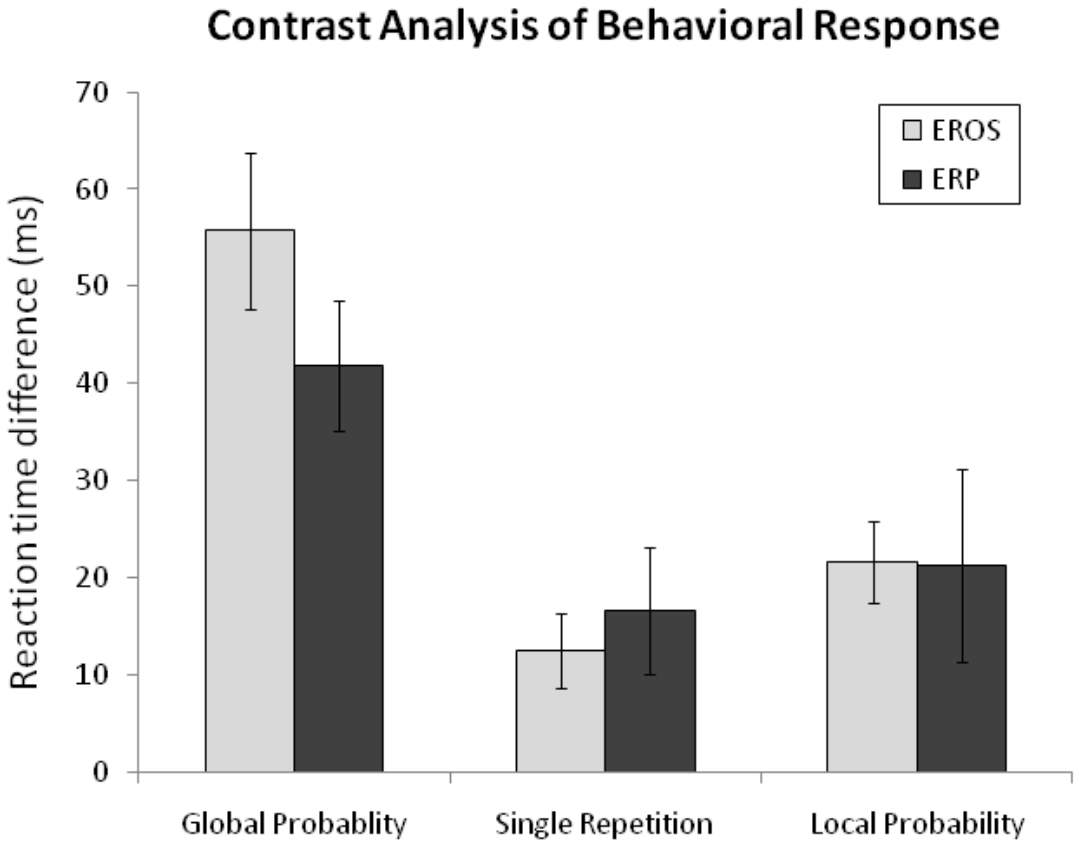


Figure 3.4. Contrast analysis of reaction time for each contingency rule and ERP/EROS session. Based on the contrast analysis described in Figure 3.1, reaction time differences (i.e. weighted sum of the reaction time of each trial type in Figure 3.3) for the Global Probability, Single Repetition, and Local Probability rules for each of the ERP and EROS recording session are presented; only data from the second ERP session and the combined EROS session are shown.

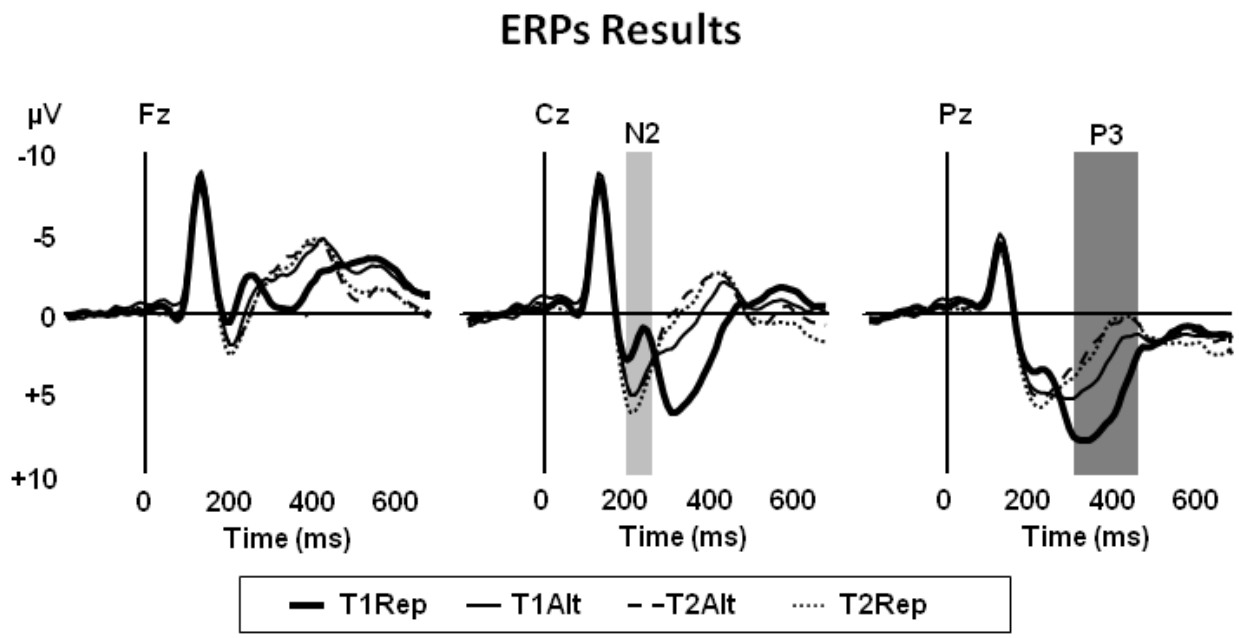


Figure 3.5. Grand average ERPs waveforms at mid line electrodes for each trial types. Differential responses to different trial types were found at 200-250ms ("N2" time window), with largest differential responses at Cz (highlighted in light gray), and 250-450ms ("P3" time window), with largest differential responses at Pz (highlighted in dark gray).

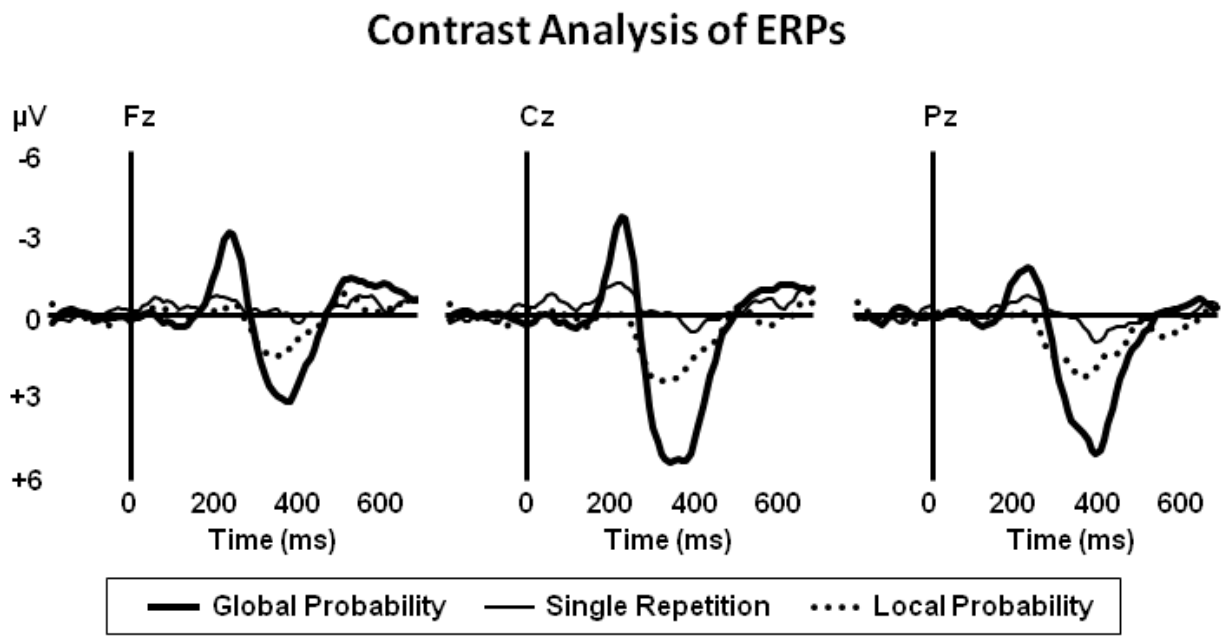


Figure 3.6. Contrast analysis of ERP waveforms for each contingency rule.

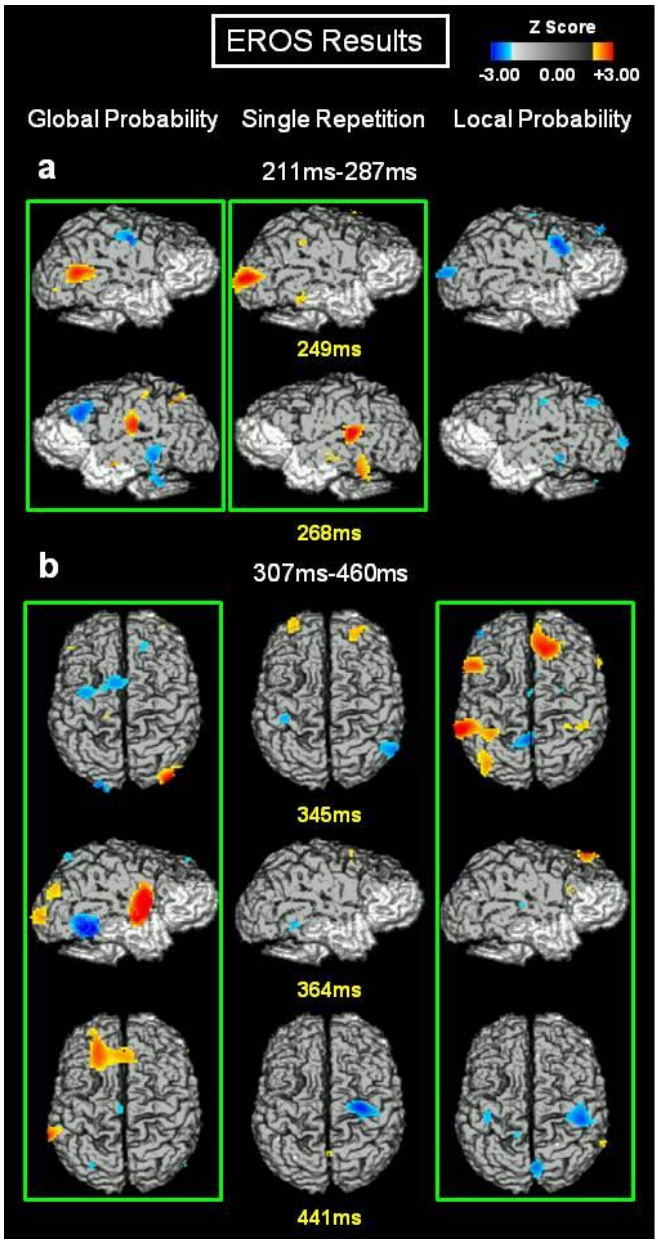


Figure 3.7. Statistical maps of EROS based on the contrasts for Global Probability (statistical), Single Repetition (relational), and Local Probability (context-related) rules. The intervals of interest are 211ms-287ms (corresponding to the timing of ERP N2 component, panel a), and 307ms-460ms (corresponding to the timing of the ERP P3 component). Only time points with statistically significant EROS responses are shown. Each time point (in yellow) corresponds to the onset of a 25.6ms sampling time window. The dark gray shading shows the area of cortex interrogated by the recording montage. Bilateral temporal activations were found for the Global Probability (statistical) and Single Repetition (relational) contrasts in the early ("N2") time window, while frontal EROS responses were found for the Global (statistical) and Local (context-related) Probability contrasts in the late ("P3") time window.

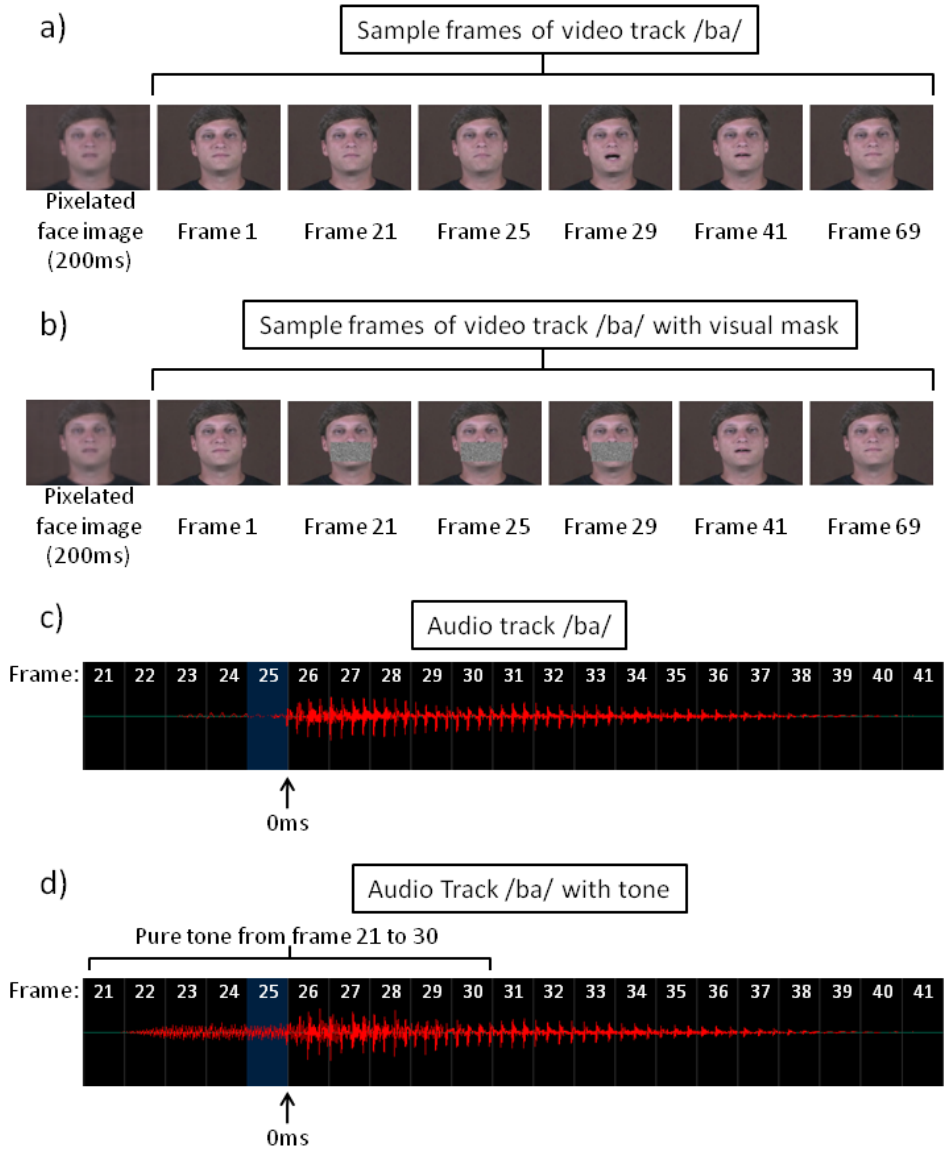


Figure 4.1. Examples of the video (panels a and b) and audio (panel c and d) tracks of the speech stimuli. a) frames sampled from the video track of /ba/. b) frames sampled from the video track of a visual catch trial with a short duration mask starting at frame 21; the duration of the gray color mask covering the mouth could be long (30 frames) or short (10 frames), and the mask could start at frame 6, 21, or 36; The video tracks were preceded by a pixelated still image of the face to ensure smooth transition between videos; c) section of the audio track /ba/ corresponding to frames 21 to 41. The end of frame 25 demarcated the onset of the speech sound and time zero in the following analyses. d) section of the audio track of the auditory catch trial. A pure tone was inserted into the audio track from frame 21 to 30. Similar to the visual mask in a visual catch trial, the duration of the tone could be long (30 frames) or short (10 frames), and the tone could start at frame 6, 21, or 36.

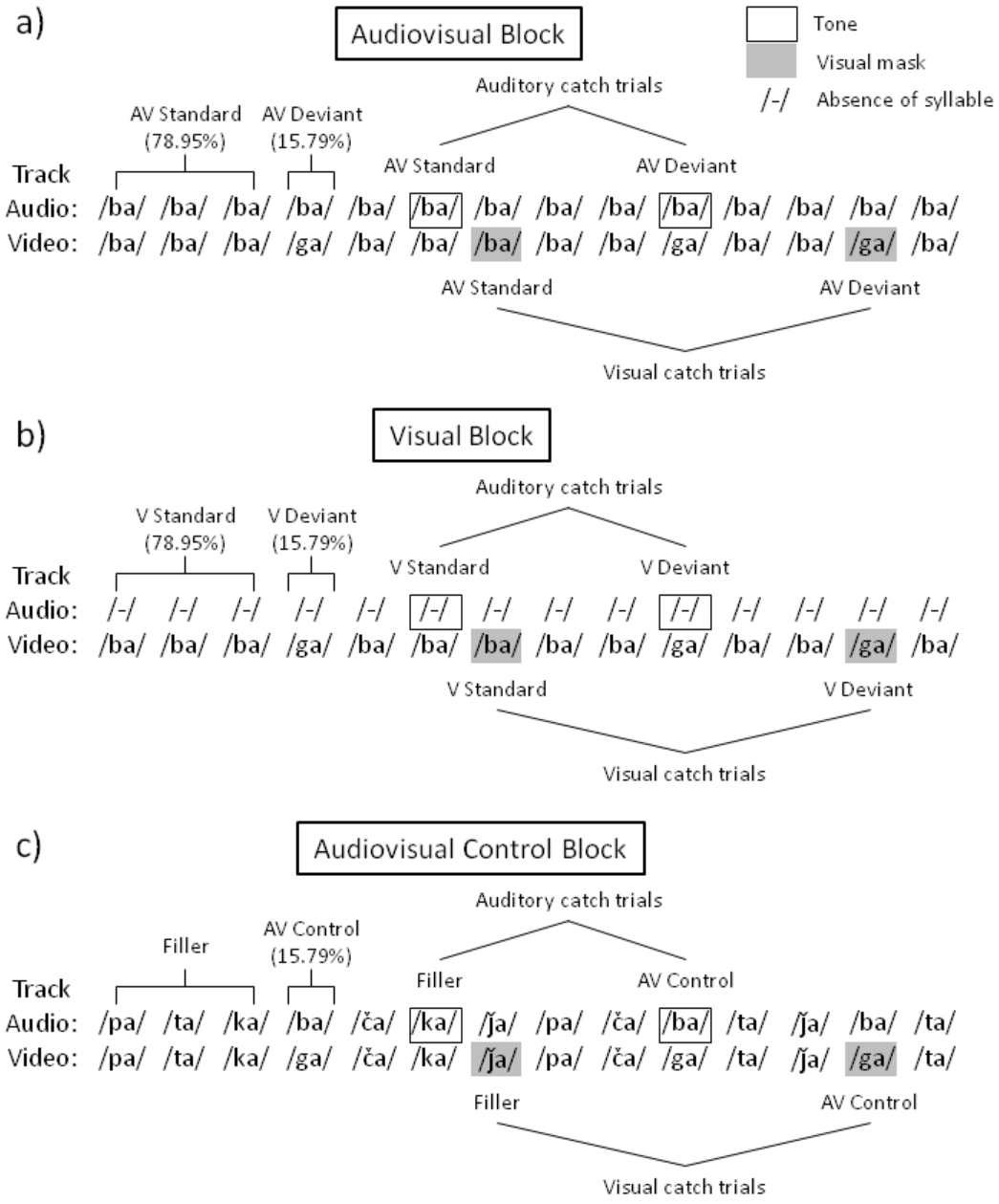


Figure 4.2. Illustration of the stimuli properties in the a) audiovisual block, b) visual block, c) audiovisual control block. In the audiovisual block 78.95% of trials were standards, A/ba/ V/ba/, 15.79% were deviants, McGurk phonemes A/ba/ V/ga/, and 5.26% were auditory or visual catch trials. The visual block was similar to the AV block except for the absence of speech sound. The AV control block consisted of six types of equally probable (15.79%) speech sound, including the McGurk stimuli. Auditory catch trials were the standard or deviant stimuli with a pure tone inserted on the sound track (empty boxes). Visual catch trials were the standard or deviant stimuli with a gray mask inserted on the video track (gray boxes). /ča/ and /ǰa/ correspond to the phonetic symbols of cha and ja, respectively.

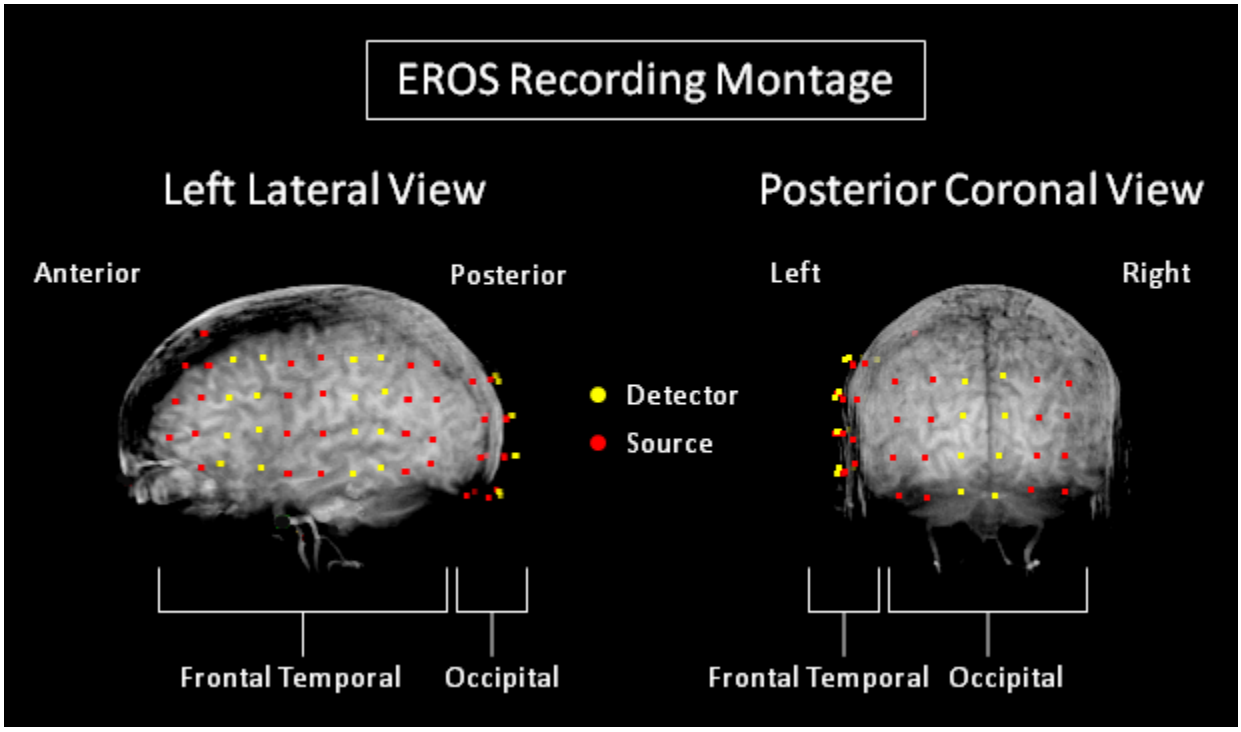


Figure 4.3. An example of the EROS recording montage overlaid on the structural MRI. The positions of the light sources (red circles) and detectors (yellow circles), covering the frontal, temporal, and occipital cortices, were projected on the left lateral (left panel) and posterior coronal (right panel) surface of the structural MRI.

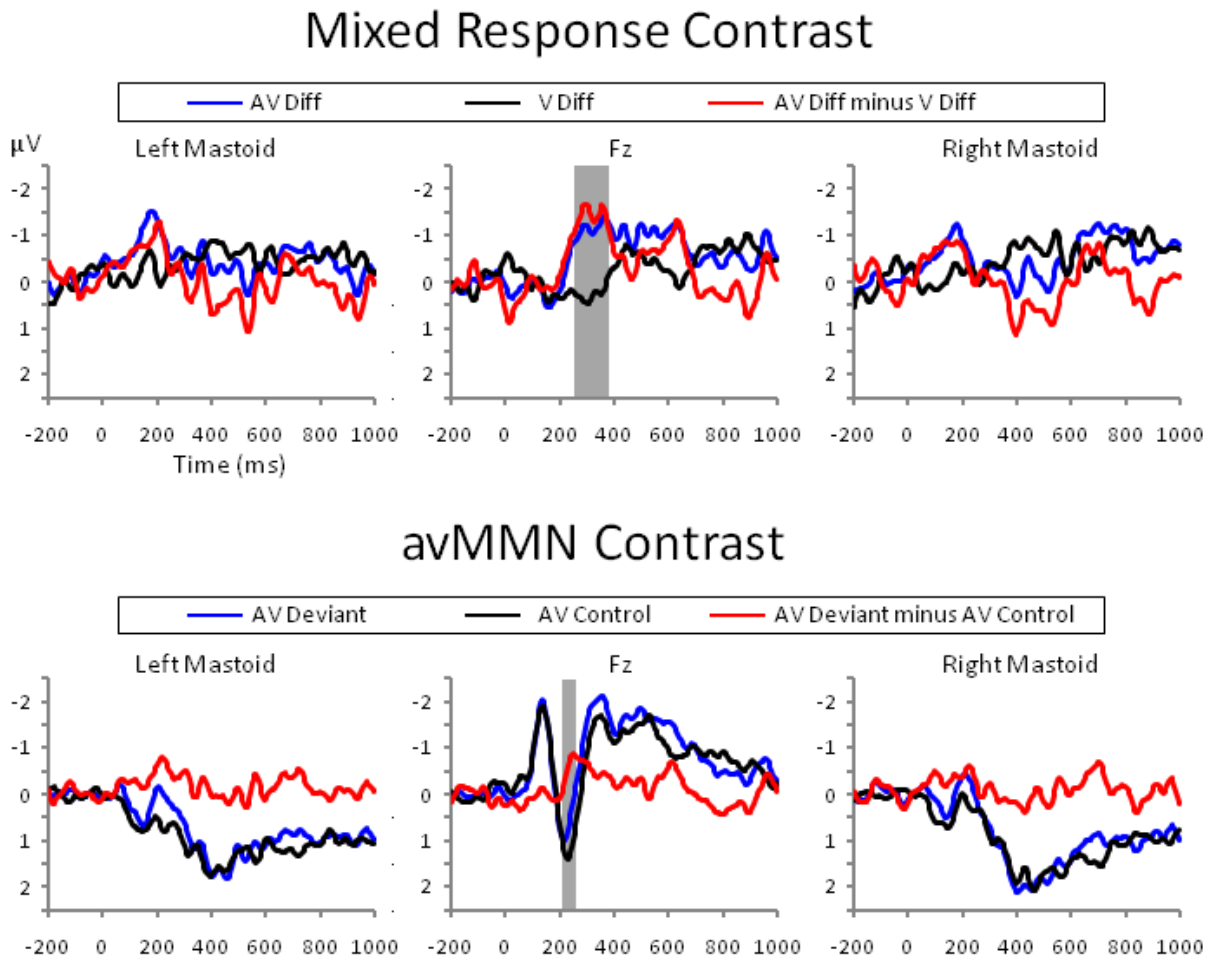


Figure 4.4. Averaged ERP waveforms for the mixed response contrast (upper panel) and avMMN contrast (lower panel) for the left mastoid (left column), Fz (middle column), and right mastoid (right column) electrodes. The AV diff waveform represents the difference waveform given by subtracting the AV standard waveform from the AVdeviant waveform, while the V diff waveform represents the difference waveform given by subtracting the V standard waveform from the V deviant waveform.

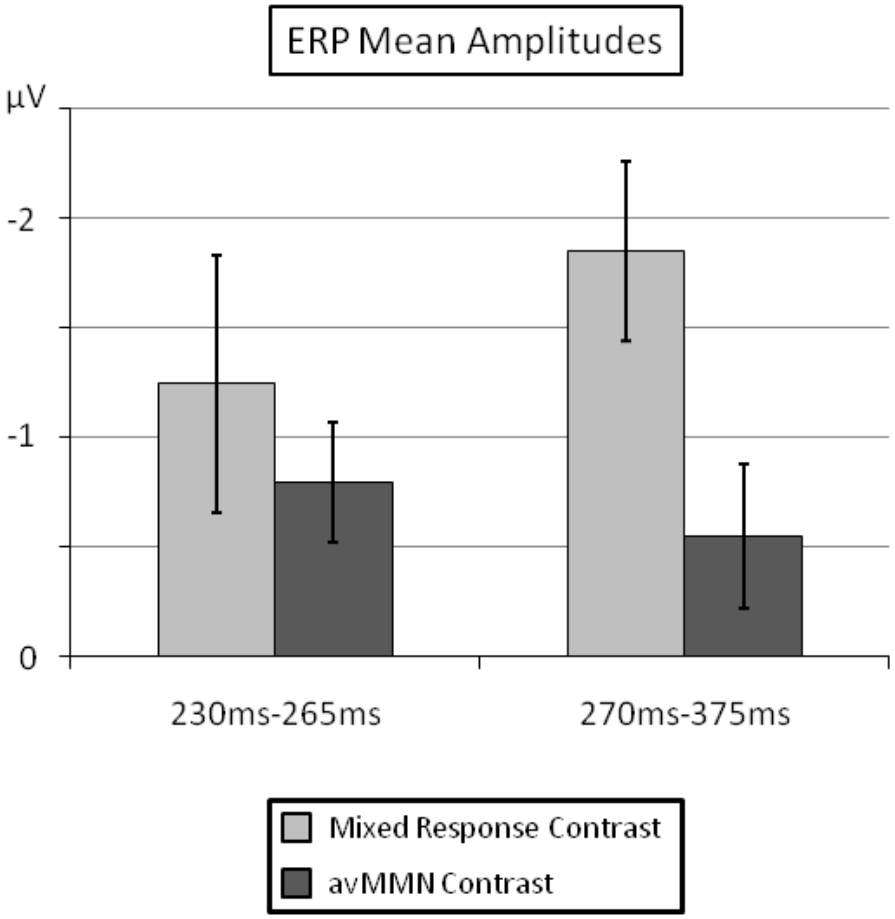


Figure 4.5. Mean ERP amplitude across participants for the early (230 ms to 265 ms) and late time windows (270 ms to 375 ms) of the mixed response and avMMN contrasts. Error bars illustrate the standard error of the mean computed across subjects.

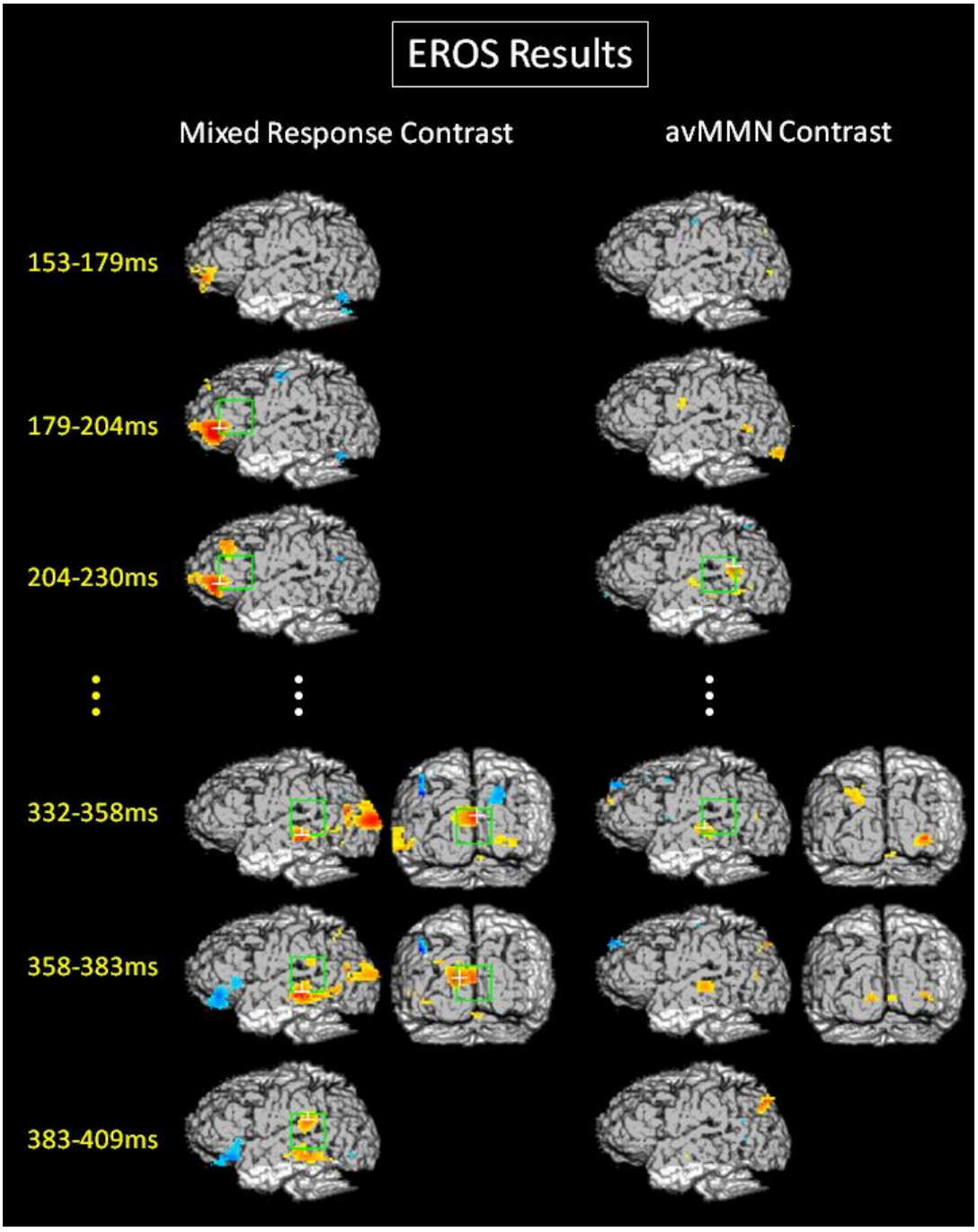


Figure 4.6. Statistical maps of EROS results overlaid on the left lateral and posterior coronal view of the template brain. The left and right columns present results from the mixed response and avMMN contrasts, respectively. The top three rows show results in the time window from 153 ms to 230 ms and the bottom three rows show results from 332 ms to 409 ms. Green boxes indicate the locations of the frontal, temporal, and occipital ROIs. The white cross within the ROI indicates the location of peak activity. The green boxes and white crosses are only shown in time frames and contrasts with significant peak responses. The dark gray shading shows the area of cortex interrogated by the recording montage.

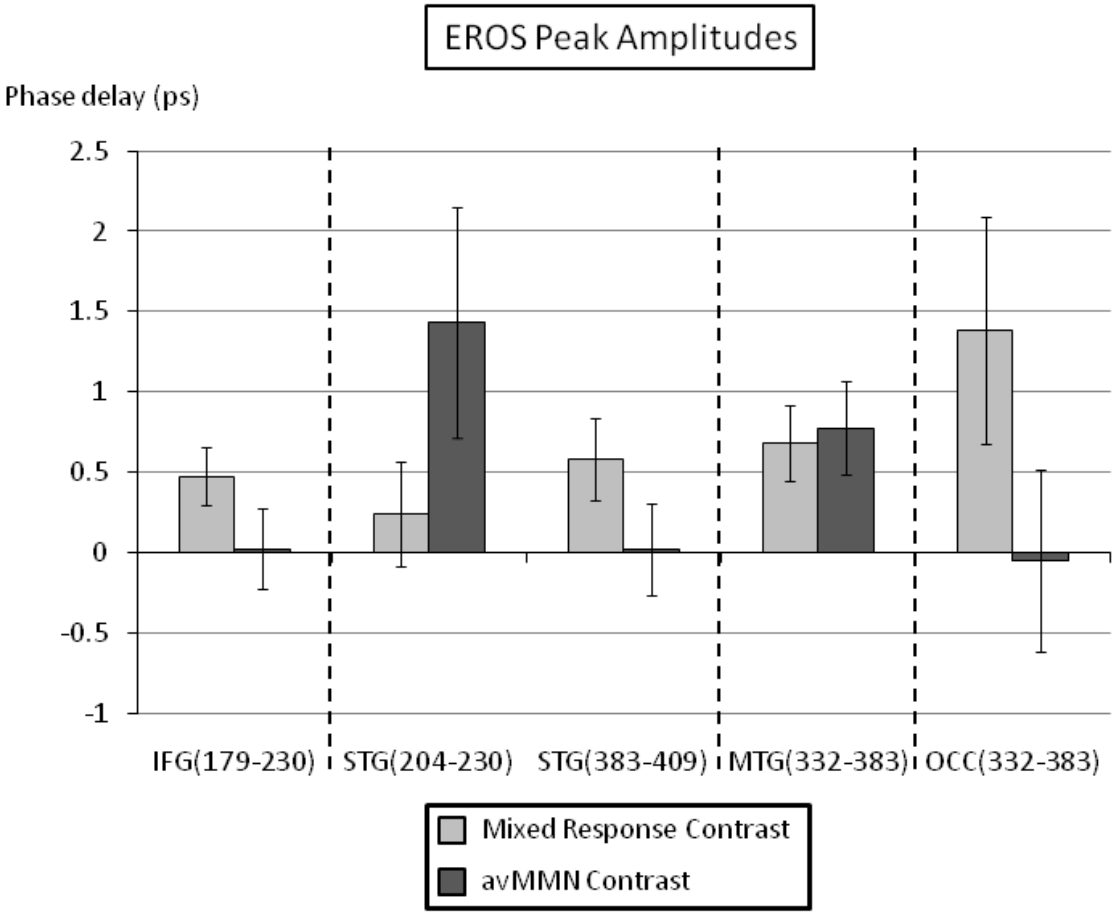


Figure 4.7. Averaged peak amplitude of the EROS response across participants for the four clusters of optical responses identified in Figure 4.6 in the mixed response and avMMN contrasts. Vertical dash lines separate the different clusters of optical response. Error bars illustrate the standard error of the mean computed across subjects.

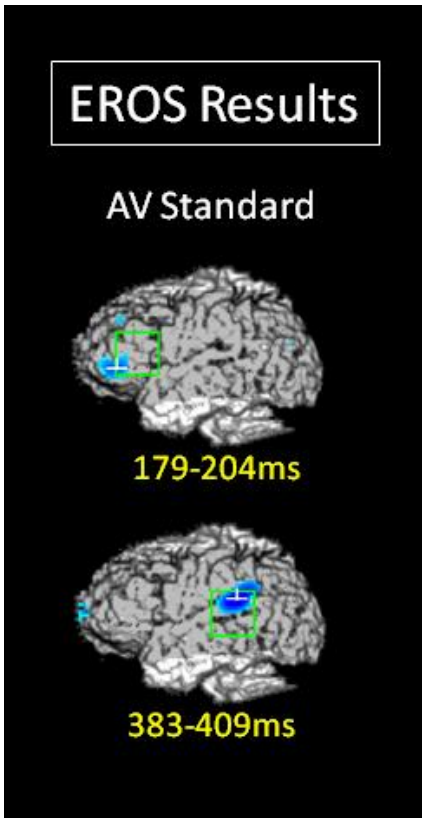


Figure 4.8. Statistical maps of EROS results for the AV standard at time windows 179 ms to 204 ms (top) and 383 ms to 409 ms (bottom). Statistically significant decreases in optical response were found in the 179 ms to 204 ms time window in the IFC (Talairach coordinate: y = 38 z = -1, peak Z = -2.63, Z critical = 2.55, $p<.05$ with correction) and in the 383 ms to 409 ms time window in the STC (y = -41 z = 17, peak Z = -2.63, Z critical = 2.55, $p<.05$ with correction).

## References


Alain, C., Hargrave, R., & Woods, D.L. (1998). Processing of auditory stimuli during visual attention in patients with schizophrenia. Biological Psychiatry, 44, 1151-1159.

Alain, C., Woods, D.L., & Knight, R.T. (1998). A distributed cortical network for auditory sensory memory in humans. Brain Research, 812, 23-37.

Alho, K., Woods, D.L., Algazi, A., Knight, R.T., & Näätänen, R. (1994). Lesions of frontal-cortex diminish the auditory mismatch negativity. Electroencephalography and Clinical Neurophysiology, 91, 353-362.

Atkinson, R.C., & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. In K.W. Spence & J.T. Spence (Eds.), The psychology of learning and motivation: Advances in research and theory (Vol. 2). New York: Academic Press.

Benar, C.G., Schon, D., Grimault, S., Nazarian, B., Burle, B., Roth, M., Badier, J.M., Marquis, P., Liegeois-Chauvel, C., & Anton, J.L. (2007). Single-trial analysis of oddball event-related potentials in simultaneous eeg-fmri. Human Brain Mapping, 28, 602-613.

Besle, J., Bertrand, O., & Giard, M.H. (2009). Electrophysiological (eeg, seeg, meg) evidence for multiple audiovisual interactions in the human auditory cortex. Hearing Research, 258, 143-151.

Besle, J., Fort, A., Delpuech, C., & Giard, M.H. (2004). Bimodal speech: Early suppressive visual effects in human auditory cortex. European Journal of Neuroscience, 20, 2225-2234.

Bischoff-Grethe, A., Goedert, K.M., Willingham, D.T., & Grafton, S.T. (2004). Neural substrates of response-based sequence learning using fmri. Journal of Cognitive Neuroscience, 16, 127-138.

Bledowski, C., Prvulovic, D., Hoechstetter, K., Scherg, M., Wibral, M., Goebel, R., & Linden, D.E.J. (2004). Localizing p300 generators in visual target and distractor processing: A combined event-related potential and functional magnetic resonance imaging study. Journal of Neuroscience, 24, 9353-9360.

Bookheimer, S. (2002). Functional mri of language: New approaches to understanding the cortical organization of semantic processing. Annual Review of Neuroscience, 25, 151-188.

Bookheimer, S.Y. (1996). Positron emission tomography studies of cognition. In E.D. Bigler (Ed.), Neuroimaging 1: Basic science. New York: Plenum Press.

Bruin, K.J., & Wijers, A.A. (2002). Inhibition, response mode, and stimulus probability: A comparative event-related potential study. Clinical Neurophysiology, 113, 1172-1182.

Brumback, C.R., Low, K.A., Gratton, G., & Fabiani, M. (2005). Putting things into perspective - individual differences in working-memory span and the integration of information. Experimental Psychology, 52, 21-30.

Budd, T.W., Barry, R.J., Gordon, E., Rennie, C., & Michie, P.T. (1998). Decrement of the n1 auditory event-related potential with stimulus repetition: Habituation vs. Refractoriness. International Journal of Psychophysiology, 31, 51-68.

Bulkin, D.A., & Groh, J.M. (2006). Seeing sounds: Visual and auditory interactions in the brain. Current Opinion in Neurobiology, 16, 415-419.

Bushell, C.M. (1996). Dissociated identity and semantic priming in broca's aphasia: How controlled processing produces inhibitory semantic priming. Brain and Language, 55, 264-288.

Calhoun, V.D., Adali, T., Pearlson, G.D., & Kiehl, K.A. (2006). Neuronal chronometry of target detection: Fusion of hemodynamic and event-related potential data. Neuroimage, 30, 544-553.

Callan, D.E., Jones, J.A., Munhall, K., Callan, A.M., Kroos, C., & Vatikiotis-Bateson, E. (2003). Neural processes underlying perceptual enhancement by visual speech gestures. Neuroreport, 14, 2213-2218.

Calvert, G.A. (2001). Crossmodal processing in the human brain: Insights from functional neuroimaging studies. Cerebral Cortex, 11, 1110-1123.

Calvert, G.A., Bullmore, E.T., Brammer, M.J., Campbell, R., Williams, S.C.R., McGuire, P.K., Woodruff, P.W.R., Iversen, S.D., & David, A.S. (1997). Activation of auditory cortex during silent lipreading. Science, 276, 593-596.

Caplan, D., Alpert, N., & Waters, G. (1998). Effects of syntactic structure and propositional number on patterns of regional cerebral blood flow. Journal of Cognitive Neuroscience, 10, 541-552.

Caramazza, A., Capitani, E., Rey, A., & Berndt, R.S. (2001). Agrammatic broca's aphasia is not associated with a single pattern of comprehension performance. Brain and Language, 76, 158-184.

Cleeremans, A., & McClelland, J.L. (1991). Learning the structure of event sequences. Journal of Experimental Psychology-General, 120, 235-253.

Cohen, L.B. (1973). Changes in neuron structure during action potential propagation and synaptic transmission. Physiological Reviews, 53, 737-418.

Colin, C., Radeau, M., Soquet, A., & Deltenre, P. (2004). Generalization of the generation of an mmn by illusory mcgurk percepts: Voiceless consonants. Clinical Neurophysiology, 115, 1989-2000.

Colin, C., Radeau, M., Soquet, A., Demolin, D., Colin, F., & Deltenre, P. (2002). Mismatch negativity evoked by the mcgurk-macdonald effect: A phonetic representation within short-term memory. Clinical Neurophysiology, 113, 495-506.

Coulson, S., King, J.W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. Language and Cognitive Processes, 13, 21-58.

Czigler, I. (2007). Visual mismatch negativity - violation of nonattended environmental regularities. Journal of Psychophysiology, 21, 224-230.

D'Arcy, R.C.N., Connolly, J.F., Service, E., Hawco, C.S., & Houlihan, M.E. (2004). Separating phonological and semantic processing in auditory sentence processing: A high-resolution event-related brain potential study. Human Brain Mapping, 22, 40-51.

Deouell, L.Y. (2007). The frontal generator of the mismatch negativity revisited. Journal of Psychophysiology, 21, 188-203.

Doeller, C.F., Opitz, B., Mecklinger, A., Krick, C., Reith, W., & Schröger, E. (2003). Prefrontal cortex involvement in preattentive auditory deviance detection: Neuroimaging and electrophysiological evidence. Neuroimage, 20, 1270-1282.

Donchin, E. (1981). Society for psychophysiological research - presidential-address, 1980 - surprise ... Surprise. Psychophysiology, 18, 493-513.

Donchin, E., & Coles, M.G.H. (1988). Is the p300 component a manifestation of context updating. Behavioral and Brain Sciences, 11, 357-374.

Duncan-Johnson, C.C., & Donchin, E. (1977). Quantifying surprise - variation of event-related potentials with subjective-probability. Psychophysiology, 14, 456-467.

Duncan-Johnson, C.C., & Donchin, E. (1982). The p300 component of the event-related brain potential as an index of information-processing. Biological Psychology, 14, 1-52.

Elman, J.L. (1990). Finding structure in time. Cognitive Science, 14, 179-211.

Fabiani, M. (2006). Multiple electrophysiological indices of distinctiveness. In R.R. Hunt & J.B. Worthen (Eds.), Distinctiveness and memory Cambridge, MA: Oxford University Press.

Fabiani, M., & Friedman, D. (1995). Changes in brain activity patterns in aging - the novelty oddball. Psychophysiology, 32, 579-594.

Fabiani, M., Gratton, G., & Federmeir, K.D. (2006). Event-related brain potentials: Methods, theory, and application. In J. Cacioppo, L. Tassinary, & G. Berntson (Eds.), Handbook of psychophysiology (pp. 85-119): Cambridge University Press.

Fabiani, M., Low, K.A., Wee, E., Sable, J.J., & Gratton, G. (2006). Reduced suppression or labile memory? Mechanisms of inefficient filtering of irrelevant information in older adults. Journal of Cognitive Neuroscience, 18, 637-650.

Fadiga, L., Craighero, L., & Olivier, E. (2005). Human motor cortex excitability during the perception of others' action. Current Opinion in Neurobiology, 15, 213-218.

Federmeier, K.D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. Psychophysiology, 44, 491-505.

Folstein, J.R., & VanPetten, C. (2008). Influence of cognitive control and mismatch on the n2 component of the erp: A review. Psychophysiology, 45, 152-170.

Franceschini, M.A., & Boas, D.A. (2004). Noninvasive measurement of neuronal activity with near-infrared optical imaging. Neuroimage, 21, 372-386.

Friederici, A.D. (2002). Towards a neural basis of auditory sentence processing. Trends in Cognitive Sciences, 6, 78-84.

Friederici, A.D., Ruschemeyer, S.A., Hahne, A., & Fiebach, C.J. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: Localizing syntactic and semantic processes. Cerebral Cortex, 13, 170-177.

Frishkoff, G.A., Tucker, D.M., Davey, C., & Scherg, M. (2004). Frontal and posterior sources of event-related potentials in semantic comprehension. Cognitive Brain Research, 20, 329-354.

Fuster, J.M. (2001). The prefrontal cortex--an update: Time is of the essence. Neuron, 30, 319-333.

Fuster, J.M. (2004). Upper processing stages of the perception-action cycle. Trends in Cognitive Sciences, 8, 143-145.

Fuster, J.M. (2006). The cognit: A network model of cortical representation. Models and Theories of Brain Function with Special Emphasis on Cognitive Processing (Vol. 60, pp. 125-132).

Fuster, J.M., Bodner, M., & Kroger, J.K. (2000). Cross-modal and cross-temporal association in neurons of frontal cortex. Nature, 405, 347-351.

Giard, M.H., Perrin, F., Pernier, J., & Bouchet, P. (1990). Brain generators implicated in the processing of auditory stimulus deviance - a topographic event-related potential study. Psychophysiology, 27, 627-640.

Goodglass, H. (1993). Understanding aphasia. San Diego: Academic Press.

Grafton, S.T., Hazeltine, E., & Ivry, R. (1995). Functional mapping of sequence learning in normal humans. Journal of Cognitive Neuroscience, 7, 497-510.

Gratton, G. (2000). "Opt-cont" And "Opt-3d": A software suite for the analysis and 3d reconstruction of the event-related optical signal (eros). Psychophysiology, 37, S44-S44.

Gratton, G., Brumback, C.R., Gordon, B.A., Pearson, M.A., Low, K.A., & Fabiani, M. (2006). Effects of measurement method, wavelength, and source-detector distance on the fast optical signal. Neuroimage, 32, 1576-1590.

Gratton, G., Coles, M.G.H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. Electroencephalography and Clinical Neurophysiology, 55, 468-484.

Gratton, G., & Corballis, P.M. (1995). Removing the heart from the brain - compensation for the pulse artifact in the photon migration signal. Psychophysiology, 32, 292-299.

Gratton, G., Corballis, P.M., Cho, E., Fabiani, M., & Hood, D.C. (1995). Shades of gray matter: Noninvasive optical images of human brain responses during visual stimulation. Psychophysiology, 32, 505-509.

Gratton, G., & Fabiani, M. (2007). Optical imaging In R. Parasuraman & M. Rizzo (Eds.), Neuroergonomics: The brain at work (pp. 65-81). Cambridge, MA: Oxford University Press.

Gratton, G., Fabiani, M., Elbert, T., & Rockstroh, B. (2003). Seeing right through you: Applications of optical imaging to the study of the human brain. Psychophysiology, 40, 487-491.

Gratton, G., Fabiani, M., Goodman-Wood, M.R., & Desoto, M.C. (1998). Memory-driven processing in human medial occipital cortex: An event-related optical signal (eros) study. Psychophysiology, 35, 348-351.

Gratton, G., Rykhlevskaia, E., Wee, E., Leaver, E., & Fabiani, M. (2009). Does white matter matter? Spatiotemporal dynamics of task switching in aging. Journal of Cognitive Neuroscience, 21, 1380-1395.

Gratton, G., Sarno, A., Maclin, E., Corballis, P.M., & Fabiani, M. (2000). Toward noninvasive 3-d imaging of the time course of cortical activity: Investigation of the depth of the event-related optical signal. Neuroimage, 11, 491-504.

Guiraud, J., Besle, J., Arnold, L., Boyle, P., Giard, M.-H., Bertrand, O., Norena, A., Truy, E., & Collet, L. (2007). Evidence of a tonotopic organization of the auditory cortex in cochlear implant users. Journal of Neuroscience, 27, 7838-7846.

Gupta, P., & Cohen, N.J. (2002). Theoretical and computational analysis of skill learning, repetition priming, and procedural memory. Psychological Review, 109, 401-448.

Hackett, T.A., Stepniewska, I., & Kaas, J.H. (1999). Prefrontal connections of the parabelt auditory cortex in macaque monkeys. Brain Research, 817, 45-58.

Hagoort, P. (2005). On broca, brain, and binding: A new framework. Trends in Cognitive Sciences, 9, 416-423.

Hagoort, P., & Brown, C.M. (2000). Erp effects of listening to speech compared to reading: The p600/sps to syntactic violations in spoken sentences and rapid serial visual presentation. Neuropsychologia, 38, 1531-1549.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K.M. (2004). Integration of word meaning and world knowledge in language comprehension. Science, 304, 438-441.

Halgren, E., Dhond, R.P., Christensen, N., Van Petten, C., Marinkovic, K., Lewine, J.D., & Dale, A.M. (2002). N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. Neuroimage, 17, 1101-1116.

Halgren, E., Marinkovic, K., & Chauvel, P. (1998). Generators of the late cognitive potentials in auditory and visual oddball tasks. Electroencephalography and Clinical Neurophysiology, 106, 156-164.

Horvath, J., Czigler, I., Sussman, E., & Winkler, I. (2001). Simultaneously active pre-attentive representations of local and global rules for sound sequences in the human brain. Cognitive Brain Research, 12, 131-144.

Jääskeläinen, I.P., Pekkonen, E., Hirvonen, J., Sillanaukee, P., & Naatanen, R. (1996). Mismatch negativity subcomponents and ethyl alcohol. Biological Psychology, 43, 13-25.

Jacobsen, T., & Schröger, E. (2001). Is there pre-attentive memory-based comparison of pitch. Psychophysiology, 38, 1-5.

January, D., Trueswell, J.C., & Thompson-Schill, S.L. (2009). Co-localization of stroop and syntactic ambiguity resolution in broca's area: Implications for the neural basis of sentence processing. Journal of Cognitive Neuroscience, 21, 2434-2444.

Just, M.A., Carpenter, P.A., Keller, T.A., Eddy, W.F., & Thulborn, K.R. (1996). Brain activation modulated by sentence comprehension. Science, 274, 114-116.

Kaan, E., & Swaab, T.Y. (2002). The brain circuitry of syntactic comprehension. Trends in Cognitive Sciences, 6, 350-356.

Kaas, J.H., & Hackett, T.A. (2000). Subdivisions of auditory cortex and processing streams in primates. Proceedings of the National Academy of Sciences of the United States of America, 97, 11793-11799.

Karis, D., Fabiani, M., & Donchin, E. (1984). "P300" And memory: Individual differehces in the von restorff effect. Cognitive Psychology, 16, 177-216.

Kekoni, J., Hämäläinen, H., Saarinen, M., Gröhn, J., Reinikainen, K., Lehtokoski, A., & Näätänen, R. (1997). Rate effect and mismatch responses in the somatosensory system: Erp-recordings in humans. Biological Psychology, 46, 125-142.

Kiehl, K.A., Laurens, K.R., & Liddle, P.F. (2002). Reading anomalous sentences: An event-related fmri study of semantic processing. Neuroimage, 17, 842-850.

Kimura, M., Katayama, J., & Murohashi, H. (2005). Positive difference in erps reflects independent processing of visual changes. Psychophysiology, 42, 369-379.

Kimura, M., Katayama, J.i., Ohira, H., & Schröger, E. (2009). Visual mismatch negativity: New evidence from the equiprobable paradigm. Psychophysiology, 46, 402-409.

Kirino, E., Belger, A., Goldman-Rakic, P., & McCarthy, G. (2000). Prefrontal activation evoked by infrequent target and novel stimuli in a visual target detection task: An event-related functional magnetic resonance imaging study. Journal of Neuroscience, 20, 6612-6618.

Knight, R.T. (1984). Decreased response to novel stimuli after prefrontal lesions in man. Electroencephalography and Clinical Neurophysiology, 59, 9-20.

Knight, R.T., Scabini, D., Woods, D.L., & Clayworth, C.C. (1989). Contributions of temporal-parietal junction to the human auditory p3. Brain Research, 502, 109-116.

Koch, I., & Hoffmann, J. (2000). Patterns, chunks, and hierarchies in serial reaction-time tasks. Psychological Research-Psychologische Forschung, 63, 22-35.

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. Science, 302, 1181-1185.

Kuperberg, G.R., McGuire, P.K., Bullmore, E.T., Brammer, M.J., Rabe-Hesketh, S., Wright, I.C., Lythgoe, D.J., Williams, S.C.R., & David, A.S. (2000). Common and distinct neural substrates for pragmatic, semantic, and syntactic processing of spoken sentences: An fmri study. Journal of Cognitive Neuroscience, 12, 321-341.

Kutas, M., & Federmeier, K.D. (2000). Electrophysiology reveals semantic memory use in language comprehension. Trends in Cognitive Sciences, 4, 463-470.

Kutas, M., & Hillyard, S. (1984). Brain potentials during reading reflect word expectancy and semantic association. Nature, 307, 161-163.

Kutas, M., & Hillyard, S.A. (1980a). Reading senseless sentences - brain potentials reflect semantic incongruity. Science, 207, 203-205.

Kutas, M., & Hillyard, S.A. (1980b). Reading senseless sentences: Brain potentials reflect semantic incongruity. Science, 207.

Kutas, M., & Hillyard, S.A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. Memory & Cognition, 11, 539-550.

Linden, D.E.J. (2005). The p300: Where in the brain is it produced and what does it tell us? Neuroscientist, 11, 563-576.

Low, K.A., Leaver, E., Kramer, A.F., Fabiani, M., & Gratton, G. (2006). Fast optical imaging of frontal cortex during active and passive oddball tasks. Psychophysiology, 43, 127-136.

Maclin, E.L., Low, K.A., Sable, J.J., Fabiani, M., & Gratton, G. (2004). The event-related optical signal to electrical stimulation of the median nerve. Neuroimage, 21, 1798-1804.

Mantini, D., Corbetta, M., Perrucci, M.G., Romani, G.L., & Del Gratta, C. (2009). Large-scale brain networks account for sustained and transient activity during target detection. Neuroimage, 44, 265-274.

McCarthy, G., Nobre, A., Bentin, S., & Spencer, D. (1995). Language-related field potentials in the anterior-medial temporal lobe: I. Intracranial distribution and neural generators. Journal of Neuroscience, 15, 1080-1089.

McGurk, H., & Macdonald, J. (1976). Hearing lips and seeing voices. Nature, 264, 746-748.

Medvedev, A.V., Kainerstorfer, J., Borisov, S.V., Barbour, R.L., & VanMeter, J. (2008). Event-related fast optical signal in a rapid object recognition task: Improving detection by the independent component analysis. Brain Research, 1236, 145-158.

Meyer, M., Friederici, A.D., & von Cramon, D.Y. (2000). Neurocognition of auditory sentence comprehension: Event related fmri reveals sensitivity to syntactic violations and task demands. Cognitive Brain Research, 9, 19-33.

Miller, L.M., & D'Esposito, M. (2005). Perceptual fusion and stimulus coincidence in the cross-modal integration of speech. Journal of Neuroscience, 25, 5884-5893.

Molholm, S., Martinez, A., Ritter, W., Javitt, D.C., & Foxe, J.J. (2005). The neural circuitry of pre-attentive auditory change-detection: An fmri study of pitch and duration mismatch negativity generators. Cerebral Cortex, 15, 545-551.

Moro, A., Tettamanti, M., Perani, D., Donati, C., Cappa, S.F., & Fazio, F. (2001). Syntax and the brain: Disentangling grammar by selective anomalies. Neuroimage, 13, 110-118.

Mottonen, R., Krause, C.M., Tiippana, K., & Sams, M. (2002). Processing of changes in visual speech in the human auditory cortex. Cognitive Brain Research, 13, 417-425.

Näätänen, R. (1992). Attention and brain function. Hillsdale, NJ: Erlbaum.

Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (mmn) and its magnetic equivalent (mmnm). Psychophysiology, 38, 1-21.

Näätänen, R. (2007). The mismatch negativity - where is the big fish? Journal of Psychophysiology, 21, 133-137.

Näätänen, R., Jacobsen, T., & Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (mmn): A review of the evidence. Psychophysiology, 42, 25-32.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R.J., Luuk, A., Allik, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. Nature, 385, 432-434.

Näätänen, R., & Michie, P.T. (1979). Early selective-attention effects on the evoked potential: A critical review and reinterpretation. Biological Psychology, 8, 81-136.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (mmn) in basic research of central auditory processing: A review. Clinical Neurophysiology, 118, 2544-2590.

Näätänen, R., & Picton, T. (1987). The n1 wave of the human electric and magnetic response to sound - a review and an analysis of the component structure. Psychophysiology, 24, 375-425.

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). 'primitive intelligence' in the auditory cortex. Trends in Neurosciences, 24, 283-288.

Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. Psychological Bulletin, 125, 826-859.

Nelson, J.K., Reuter-Lorenz, P.A., Persson, J., Sylvester, C.-Y.C., & Jonides, J. (2009). Mapping interference resolution across task domains: A shared control process in left inferior frontal gyrus. Brain Research, 1256, 92-100.

Newman, A.J., Pancheva, R., Ozawa, K., Neville, H.J., & Ullman, M.T. (2001). An event-related fmri study of syntactic and semantic violations. Journal of Psycholinguistic Research, 30, 339-364.

Newman, S.D., Just, M.A., Keller, T.A., Roth, J., & Carpenter, P.A. (2003). Differential effects of syntactic and semantic processing on the subregions of broca's area. Cognitive Brain Research, 16, 297-307.

Ni, W., Constable, R.T., Mencl, W.E., Pugh, K.R., Fulbright, R.K., Shaywitz, S.E., Shaywitz, B.A., Gore, J.C., & Shankweiler, D. (2000). An event-related neuroimaging study distinguishing form and content in sentence processing. Journal of Cognitive Neuroscience, 12, 120-133.

Nomura, E., Maddox, W., Filoteo, J., Ing, A., Gitelman, D., Parrish, T., Mesulam, M.-M., & Reber, P. (2007). Neural correlates of rule-based and information-integration visual category learning. Cerebral Cortex, 17, 37-43.

Oldfield, R.C. (1971). The assessment and analysis of handedness: The edinburgh inventory. Neuropsychologia, 9, 97-113.

Opitz, B., & Friederici, A.D. (2004). Brain correlates of language learning: The neuronal dissociation of rule-based versus similarity-based learning. Journal of Neuroscience, 24, 8436-8440.

Opitz, B., Rinne, T., Mecklinger, A., Cramon, D.Y.v., & Schröger, E. (2002). Differential contribution of frontal and temporal cortices to auditory change detection: Fmri and erp results. Neuroimage, 15, 167-174.

Osterhout, L. (1997). On the brain response to syntactic anomalies: Manipulations of word position and word class reveal individual differences. Brain and Language, 59, 494-522.

Osterhout, L., & Holcomb, P.J. (1992). Event-related brain potentials elicited by syntactic anomaly. Journal of Memory and Language, 31, 785-806.

Paavilainen, P., Jiang, D.M., Lavikainen, J., & Naatanen, R. (1993). Stimulus-duration and the sensory memory trace - an event-related potential study. Biological Psychology, 35, 139-152.

Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G., & Elbert, T. (1995). Specific tonotopic organizations of different areas of the human auditory-cortex revealed by simultaneous magnetic and electric recordings. Electroencephalography and Clinical Neurophysiology, 94, 26-40.

Parker, G.J.M., Luzzi, S., Alexander, D.C., Wheeler-Kingshott, C.A.M., Clecarelli, O., & Ralph, M.A.L. (2005). Lateralization of ventral and dorsal auditory-language pathways in the human brain. Neuroimage, 24, 656-666.

Petersson, K.M., Forkstam, C., & Ingvar, M. (2004). Artificial syntactic violations activate broca's region. Cognitive Science, 28, 383-407.

Picton, T.W., Alain, C., Otten, L., Ritter, W., & Achim, A. (2000). Mismatch negativity: Different water in the same river. Audiology and Neuro-Otology, 5, 111-139.

Polich, J. (2007). Updating p300: An integrative theory of p3a and p3b. Clinical Neurophysiology, 118, 2128-2148.

Poline, J.B., Worsley, K.J., Evans, A.C., & Friston, K.J. (1997). Combining spatial extent and peak intensity to test for activations in functional imaging. Neuroimage, 5, 83-96.

Pritchard, W.S. (1981). Psychophysiology of p300. Psychological Bulletin, 89, 506-540.

Pulvermuller, F., & Shtyrov, Y. (2006). Language outside the focus of attention: The mismatch negativity. As a tool for studying higher cognitive processes. Progress in Neurobiology, 79, 49-71.

Rector, D.M., Carter, K.M., Volegov, P.L., & George, J.S. (2005). Spatio-temporal mapping of rat whisker barrels with fast scattered light signals. Neuroimage, 26, 619-627.

Rinne, T., Alho, K., Ilmoniemi, R.J., Virtanen, J., & Näätänen, R. (2000). Separate time behaviors of the temporal and frontal mismatch negativity sources. Neuroimage, 12, 14-19.

Rinne, T., Degerman, A., & Alho, K. (2005). Superior temporal and inferior frontal cortices are activated by infrequent sound duration decrements: An fmri study. Neuroimage, 26, 66-72.

Rinne, T., Gratton, G., Fabiani, M., Cowan, N., Maclin, E., Stinard, A., Sinkkonen, J., Alho, K., & Näätänen, R. (1999). Scalp-recorded optical signals make sound processing in the auditory cortex visible. Neuroimage, 10, 620-624.

Rockland, K.S., & Ojima, H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey. International Journal of Psychophysiology, 50, 19-26.

Ross, L.A., Saint-Amour, D., Leavitt, V.M., Javitt, D.C., & Foxe, J.J. (2007). Do you see what i am saying? Exploring visual enhancement of speech comprehension in noisy environments. Cerebral Cortex, 17, 1147-1153.

Ruschemeyer, S.A., Fiebach, C.J., Kempe, V., & Friederici, A.D. (2005). Processing lexical semantic and syntactic information in first and second language: Fmri evidence from german and russian. Human Brain Mapping, 25, 266-286.

Rykhlevskaia, E., Gratton, G., & Fabiani, M. (2008). Combining structural and functional neuroimaging data for studying brain connectivity: A review. Psychophysiology, 45, 173-187.

Saarinen, J., Paavilainen, P., Schoger, E., Tervaniemi, M., & Naatanen, R. (1992). Representation of abstract attributes of auditory-stimuli in the human brain. Neuroreport, 3, 1149-1151.

Sable, J.J., Low, K.A., Maclin, E.L., Fabiani, M., & Gratton, G. (2004). Latent inhibition mediates n1 attenuation to repeating sounds. Psychophysiology, 41, 636-642.

Sable, J.J., Low, K.A., Whalen, C., Maclin, E.L., Fabiani, M., & Gratton., G. (2004). Dissociation of mismatch responses to sound interval in human auditory cortex using noninvasive optical imaging., Society for Neuroscience, . Washington, DC.

Sable, J.J., Low, K.A., Whalen, C.J., Maclin, E.L., Fabiani, M., & Gratton, G. (2007). Optical imaging of temporal integration in human auditory cortex. European Journal of Neuroscience, 25, 298-306.

Saint-Amour, D., De Sanctis, P., Molholm, S., Ritter, W., & Foxe, J.J. (2007). Seeing voices: High-density electrical mapping and source-analysis of the multisensory mismatch negativity evoked during the mcgurk illusion. Neuropsychologia, 45, 587-597.

Sallinen, M., & Lyytinen, H. (1997). Mismatch negativity during objective and subjective sleepiness. Psychophysiology, 34, 694-702.

Sams, M., Aulanko, R., Hämäläinen, M., Hari, R., Lounasmaa, O.V., Lu, S.-T., & Simola, J. (1991). Seeing speech: Visual information from lip movements modifies activity in the human auditory cortex. Neuroscience Letters, 127, 141-145.

Sams, M., Manninen, P., Surakka, V., Helin, P., & Katto, R. (1998). Mcgurk effect in finnish syllables, isolated words, and words in sentences: Effects of word meaning and sentence context. Speech Communication, 26, 75-87.

Sams, M., Mottonen, R., & Sihvonen, T. (2005). Seeing and hearing others and oneself talk. Cognitive Brain Research, 23, 429-435.

Saur, D., Kreher, B.W., Schnell, S., Kummerer, D., Kellmeyer, P., Vry, M.S., Umarova, R., Musso, M., Glauche, V., Abel, S., Huber, W., Rijntjes, M., Hennig, J., & Weiller, C. (2008). Ventral and dorsal pathways for language. Proceedings of the National Academy of Sciences of the United States of America, 105, 18035-18040.

Schnyer, D.M., Maddox, W.T., Ell, S., Davis, S., Pacheco, C., & Verfaellie, M. (2009). Prefrontal contributions to rule-based and information-integration category learning. Neuropsychologia, 47, 2995-3006.

Schröger, E. (2007). Mismatch negativity - a microphone into auditory memory. Journal of Psychophysiology, 21, 138-146.

Shtyrov, Y., & Pulvermuller, F. (2007). Language in the mismatch negativity design. Journal of Psychophysiology, 21, 176-187.

Simos, P.G., Basile, L.F.H., & Papanicolaou, A.C. (1997). Source localization of the n400 response in a sentence-reading paradigm using evoked magnetic fields and magnetic resonance imaging. Brain Research, 762, 29-39.

Skipper, J.I., Nusbaum, H.C., & Small, S.L. (2005). Listening to talking faces: Motor cortical activation during speech perception. Neuroimage, 25, 76-89.

Skipper, J.I., van Wassenhove, V., Nusbaum, H.C., & Small, S.L. (2007). Hearing lips and seeing voices: How cortical areas supporting speech production mediate audiovisual speech perception. Cerebral Cortex, 17, 2387-2399.

Soto-Faraco, S., Navarra, J., & Alsius, A. (2004). Assessing automaticity in audiovisual speech integration: Evidence from the speeded classification task. Cognition, 92, B13-B23.

Squires, K., Petuchowski, S., Wickens, C., & Donchin, E. (1977). Effects of stimulus sequence on event related potentials - comparison of visual and auditory sequences. Perception & Psychophysics, 22, 31-40.

Sumby, W.H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. Journal of the Acoustical Society of America, 26, 212-215.

Sundara, M., Namasivayam, A.K., & Chen, R. (2001). Observation-execution matching system for speech: A magnetic stimulation study. Neuroreport, 12, 1341-1344.

Sussman, E., Ritter, W., & Vaughan, H.G. (1998). Predictability of stimulus deviance and the mismatch negativity. Neuroreport, 9, 4167-4170.

Talairach, J., & Tournoux, P. (1988). A co-planar stereotaxic atlas of a human brain. Stuttgart: Thieme.

Tervaniemi, M., Medvedev, S.V., Alho, K., Pakhomov, S.V., Roudas, M.S., Zuijen, T.L.v., & Näätänen, R. (2000). Lateralized automatic auditory processing of phonetic versus musical information: A pet study. Human Brain Mapping, 10, 74-79.

Tremblay, P., & Gracco, V.L. (2009). Contribution of the pre-sma to the production of words and non-speech oral motor gestures, as revealed by repetitive transcranial magnetic stimulation (rtms). Brain Research, 1268, 112-124.

Tse, C.-Y., Lee, C.L., Sullivan, J., Garnsey, S.M., Dell, G.S., Fabiani, M., & Gratton, G. (2007). Imaging cortical dynamics of language processing with the event-related optical signal. Proceedings of the National Academy of Sciences of the United States of America, 104, 17157-17162.

Tse, C.-Y., Ng, E., Rinne, T., & Penney, T. (2009). Temporal-frontal network underlying pre-attentive change detection: A lagged correlation path model of the event-related optical signal (eros). Frontiers in Human Neuroscience. Conference Abstract: MMN 09 Fifth Conference on Mismatch Negativity (MMN) and its Clinical and Scientific Applications, doi: 10.3389/conf.neuro.09.2009.05.097

Tse, C.-Y., & Penney, T.B. (2007). Preattentive change detection using the event-related optical signal - optical imaging of cortical activity elicited by unattended temporal deviants. IEEE Engineering in Medicine and Biology Magazine, 26, 52-58.

Tse, C.-Y., & Penney, T.B. (2008). On the functional role of temporal and frontal cortex activation in passive detection of auditory deviance. Neuroimage, 41, 1462-1470.

Tse, C.-Y., Tien, K.-R., & Penney, T.B. (2006). Event-related optical imaging reveals the temporal dynamics of right temporal and frontal cortex activation in pre-attentive change detection. Neuroimage, 29, 314-320.

van Herten, M., Kolk, H.H.J., & Chwilla, D.J. (2005). An erp study of p600 effects elicited by semantic anomalies. Cognitive Brain Research, 22, 241-255.

Van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open-class and closed-class words. Memory & Cognition, 19, 95-112.

van Wassenhove, V., Grant, K.W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. Proceedings of the National Academy of Sciences of the United States of America, 102, 1181-1186.

van Wassenhove, V., Grant, K.W., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. Neuropsychologia, 45, 598-607.

van Zuijen, T.L., Sussman, E., Winkler, I., Naatanen, R., & Tervaniemi, M. (2005). Auditory organization of sound sequences by a temporal or numerical regularity - a mismatch negativity study comparing musicians and non-musicians. Cognitive Brain Research, 23, 270-276.

Wendelken, C., Nakhabenko, D., Donohue, S.E., Carter, C.S., & Bunge, S.A. (2008). "Brain is to thought as stomach is to ??": Investigating the role of rostrolateral prefrontal cortex in relational reasoning. Journal of Cognitive Neuroscience, 20, 682-693.

Whalen, C., Maclin, E.L., Fabiani, M., & Gratton, G. (2008). Validation of a method for coregistering scalp recording locations with 3d structural mr images. Human Brain Mapping, 29, 1288-1301.

Williamson, S.J., & Kaufman, L. (1981). Biomagnetism topical review. Journal of Magnetism and Magnetic Materials, 22, 129-201.

Windmann, S. (2004). Effects of sentence context and expectation on the mcgurk illusion. Journal of Memory and Language, 50, 212-230.

Winkler, I. (2007). Interpreting the mismatch negativity. Journal of Psychophysiology, 21, 147-163.

Winkler, I., Horvath, J., Weisz, J., & Trejo, L.J. (2009). Deviance detection in congruent audiovisual speech: Evidence for implicit integrated audiovisual memory representations. Biological Psychology, 82, 281-292.

Wolf, M., Wolf, U., Choi, J.H., Gupta, R., Safonova, L.P., Paunescu, L.A., Michalos, A., & Gratton, E. (2002). Functional frequency-domain near-infrared spectroscopy detects fast neuronal signal in the motor cortex. Neuroimage, 17, 1868-1875.

Wolf, U., Wolf, M., Toronov, V., Michalos, A., Paunescu, L.A., & Gratton, E. (2000). Detecting cerebral functional slow and fast signals by frequency-domain near-infrared spectroscopy using two different sensors. OSA Meeting in Optical Spectroscopy and Imaging and Photon Migration. Miami, FL.

Yabe, H., Tervaniemi, M., Reinikainen, K., & Näätänen, R. (1997). Temporal window of integration revealed by mmn to sound omission. Neuroreport, 8, 1971-1974.

Yabe, H., Tervaniemi, M., Sinkkonen, J., Huotilainen, M., Ilmoniemi, R.J., & Näätänen, R. (1998). Temporal window of integration of auditory information in the human brain. Psychophysiology, 35, 615-619.